CHARACTERIZATION OF RNA GENETIC REGULATORS AND SYNTHETIC NETWORKS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Alexandra Michon Westbrook

August 2018

© 2018 Alexandra Michon Westbrook

ALL RIGHTS RESERVED

CHARACTERIZATION OF RNA GENETIC REGULATORS AND SYNTHETIC NETWORKS

Alexandra Michon Westbrook, Ph.D.

Cornell University 2018

A central tenent of synthetic biology is the ability to predictably engineer complex patterns of gene expression. This fined tuned control allows us to reprogram organisms with sophisticated synthetic behaviors such as producing vital chemicals and drugs and sensing environmental signals. In order to do this we need libraries of highly efficient genetic regulators and proven methods of combining them into networks. RNA presents the ideal tool to build new genetic networks because its structural and temporal characteristics allow engineers to construct fast, designable genetic networks. In this work, we show characterization and optimization of new and existing RNA regulators as well as efforts to create new behaviors with RNA-based genetic networks. We begin by vastly improving the dynamic range of an existing transcriptional RNA regulator, the pT181 attenuator, by adding translational regulation. This dual control attenuator is successfully used to reduce circuit leak in an RNA-only cascade. In order to expand upon the functionality of the RNA repressors, we design sequesters that allow us to dial down repression. The sequestration effectively creates a threshold which we use to tune the relationship between the input and output of a system. As we construct more complex circuits with diverse parts, modularity becomes essential in order to predict circuit behavior. We explore the modularity of our RNA regulators in combination with clustered regularly interspaced short palindromic repeats (CRISPR) interference (CRISPRi) in construction of an RNA pulse generator. Finally, we explore the design and implementation two complex circuits: a communication network for delivering complex signals to cells and a control network to reduce noise in biological

systems. We anticipate that the design rules learned and the tools developed here will allow construction of even more sophisticated behaviors as the growing discipline of genetic design matures.

BIOGRAPHICAL SKETCH

Alexandra was raised in Los Altos, California, by parents Anne and Scott Westbrook. She graduated from Mountain View High School in 2009 and went on to study Chemical Engineering at MIT. While at MIT, she developed a love for research in the labs of Professor Jacquin Niles in the Department of Biological Engineering and Kristala Prather in the Department of Chemical Engineering. Alexandra earned her B.S. in Chemical Engineering in June 2013. In August 2013, Alexandra entered the PhD program at Cornell's School of Chemical and Biomolecular engineering and joined the Lucks Lab later that fall.

This work is dedicated to my parents.

ACKNOWLEDGEMENTS

To my parents for their incredible support and love throughout my life. I could never have done any of this without you. Mom, thank you for your relentless drive to make sure I had every opportunity to pursue my dreams and your constant support and advice. Dad, thank you for never ignoring my 'why's when I was a kid and going through biological network and electrical circuit designs with me during my PhD. I hope we have many similar conversations to come!

To Julius, thank you so much for your guidance and support. Your love for science is infectious and always inspired me when I needed it most.

To the members of the Lucks Lab, thank you for creating a fun and collaborative environment. Katherine and Sarai, those bubble tea, chocolate, and coffee breaks were some of my best memories.

To Joel, thank you for being the most amazing partner I could ever imagine. You've always supported me and I can't imagine the past few months without you.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Engineering gene expression

The ability to predictably manipulate genes has allowed us to understand and treat diseases, program microbes to sense environmental signals, and produce useful biological molecules such as drugs and specialty chemicals (Khalil & Collins 2010). We do this by controlling gene expression using natural and synthetic genetic networks, collections of DNA elements that interact with each other through their RNA or protein expression products (Figure 1.1). Networks regulate expression of elements that make decisions about which genes to express, that in turn carry out specific functions by controlling cellular behavior. Genetic networks are built by using regulators that interact with aspects of transcription, translation, or degradation of genetic parts.

Early on, these genetic regulators were discovered in nature and then rewired to create simple networks. In order for more sophisticated networks to be built, large libraries of orthogonal regulators must constructed (Purnick & Weiss 2009). While there have been attempts to build libraries of synthetic transcription factors, proteins are notoriously difficult to design. With the discovery of zinc-fingers (Beerli & Barbas 2002) and transcription activator-like effectors (TALE) (Machens et al. 2017), engineering libraries of protein regulators has become possible, but the number of orthogonal protein parts remains limited and validating new ones is labor-intensive. Due to its smaller sequence space and more accurate folding predictions, RNA orthogonality has proven far easier to design computationally (Chappell et al. 2017; Green et al. 2014) and the discovery of CRISPRi melds RNA-designability with transcription factor-like gene expression control (Qi et al. 2013).

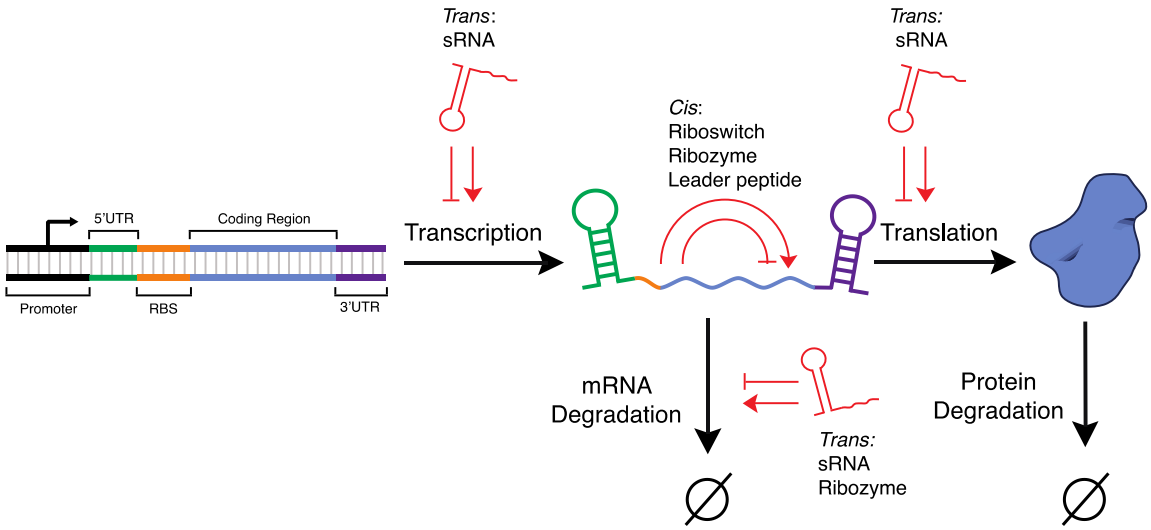


Figure 1.1 Central dogma of prokaryotic gene expression. DNA is transcribed into RNA that is in turn translated into protein. RNA and protein regulators can activate or repress gene expression by acting on transcription, translation, and degradation. Figure from Chappell et al. (Chappell et al. 2013)

1.2 The development of steady state and dynamic synthetic gene networks

Early genetic networks have focused on the ability to control the steady state ON and OFF gene expression level of a gene or multiple genes. One of the very first was a genetic toggle switch (Gardner et al. 2000) that uses two repressible promoters to remember a genetic 'state' (Figure 1.2) with bistability. Synthetic biologists have also built logic gates to control steady state behavior as a function of multiple inputs (Anderson et al. 2007; Moon et al. 2012). Recently, these logic gates have gotten quite complex, allowing for layered logic within mammalian cells (Ausländer et al. 2012), signal integration within a single mechanism (Green et al. 2017), and advanced function generation (Olson et al. 2014). Logic

gates have even been used to implement biological edge detection (Tabor et al. 2009) and pattern formation (Basu et al. 2005).

Researchers have progressed to building synthetic networks with useful dynamic properties. Dynamic activation of metabolic pathway enzymes using quorum sensing has led to increase in product titers (Gupta et al. 2017) indicating that another way to optimize circuit capabilities is through its dynamics. Constructing and optimizing circuit dynamics can help us better understand their natural counterparts. Oscillatory networks are common in nature, but the first synthetic oscillator was built in 2000 using interconnected repressors (Elowitz & Leibler 2000). This oscillator has since been optimized for tunability (Stricker et al. 2008) and further engineered for synchronous oscillations (Potvin-Trottier et al. 2016) in which it was understood that simplicity and understanding noise may be the key to precision in such dynamic networks.

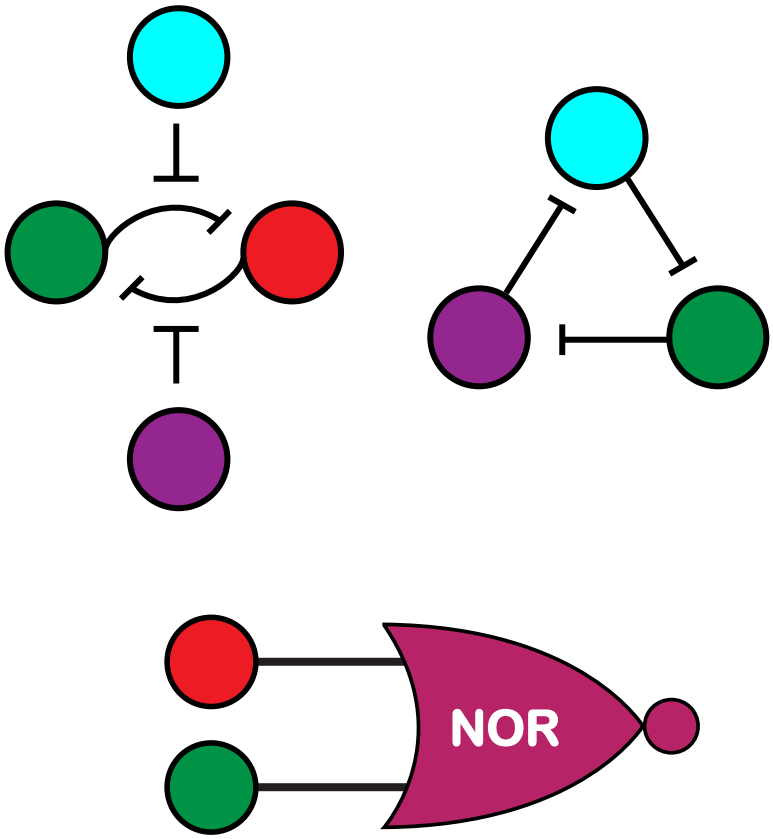


Figure 1.2 Examples of genetic networks. A genetic toggle switch (Top left), an oscillator made of connected repressors (top right), and a NOR logic gate (bottom). Circles indicate control points that can be genes or small molecules and function as inputs, out puts, or intermediates. Connecting lines represent the function a gene performs such as repression (blunted line).

1.3 RNA-based genetic regulation in bacteria

RNAs are now understood to play broad regulatory roles across the cell. As such, synthetic biologists aim to use these versatile natural systems to create a broad array of parts that can regulate many aspects of gene expression from transcription (Brantl & Wagner

2000; Qi et al. 2013), translation (Isaacs et al. 2004), and mRNA degradation (Carrier & Keasling 1999) through the conditional formation of hairpin structures at specific points in mRNAs (Chappell, Watters, et al. 2015). Transcriptional terminators repress transcription when they form by causing polymerase to ratchet off the DNA complex (Brantl & Wagner 2000), RBS-sequestering hairpins block translation by inhibiting ribosome binding (Green et al. 2014), and stability hairpins can block mRNA degradation (Carrier & Keasling 1997). RNA transcriptional regulators are particularly interesting because they regulate RNA synthesis as a function of RNA input and can be used to create genetic circuitry (Lucks et al. 2011). These circuits have many potential advantages over proteins including the possibility of leveraging RNA folding algorithms and design rules and increasing the speed of signal propagation (Takahashi et al. 2015).

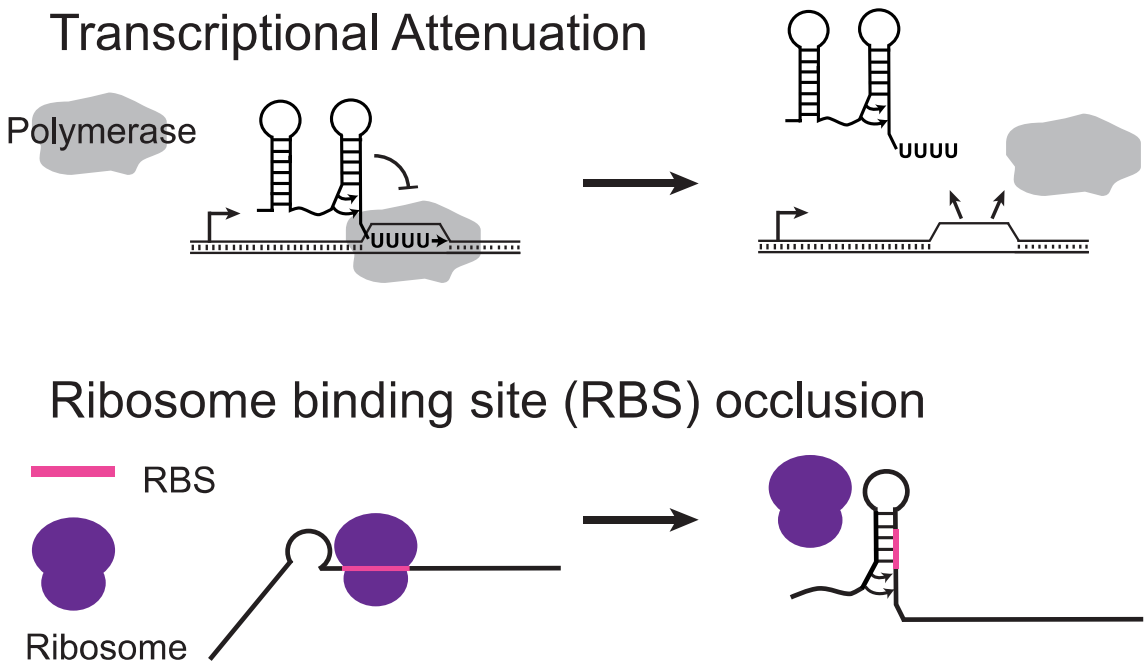


Figure 1.3 Schematic of transcriptional attenuation (top) and Ribosome binding site (RBS) occlusion (bottom).

1.3.1 Transcriptional regulation – intrinsic termination

Transcriptional regulation is often understood as acting through proteins that recruit RNA polymerase to a promoter or block transcription initiation, but RNA can also play important roles in transcriptional regulation (Qi et al. 2013; Chappell, Takahashi, et al. 2015; Liu et al. 2011). Intrinsic termination is one such mechanism of regulation in which an RNA structure called an intrinsic terminator hairpin halts transcription by causing RNA polymerase to dissociate from the DNA (Figure 1.3, top). The hairpin is usually GC-rich and followed by a poly-U sequence. The U-rich sequence causes the RNA polymerase to pause and the formation of the hairpin is thought to pull the RNA out of the polymerase and release it from the DNA. These terminator hairpins are what terminate transcription after a gene has been fully transcribed in order to release transcription machinery, but they can also conditionally form upstream of a gene as a form of regulation. The formation of the intrinsic terminator is biased by the presence of another molecule that can be another RNA in the case of the pT181 transcriptional terminator (see section 1.2.3) or by a small molecule for a transcriptional riboswitch (Watters et al. 2016).

1.3.2 Translational regulation – RBS occlusion

RNAs can also regulate one another after transcription but before the mRNA is translated into protein (Brantl 2007). Translation begins when the ribosome binds to the ribosome binding site in the 5' untranslated region (5' UTR) of the mRNA. Within the RBS is an 8 nucleotide region called the Shine-Dalgarno sequence (SD) that is partially complementary to an RNA component of the ribosome, the 16S rRNA. If RNA secondary structures block access to the SD sequence this can greatly reduce the translation rate. The formation of these

secondary structures is a mechanism of translational regulation is called ribosome binding site (RBS) occlusion (Figure 1.3, bottom) (Espah Borujeni et al. 2014). RBS accessibility can be governed by the presence of a small molecule such as for riboswitches (Breaker 2012), the formation of an RNA structure like a ribozyme (Tang & Breaker 1997), or the presence of another RNA for synthetic toehold switches (Green et al. 2014) and other natural regulators (Kittle et al. 1989; Morfeldt et al. 1995).

1.3.3 Using pT181 to build RNA-based networks

This work relies heavily on the pT181 transcriptional attenuator from the Staphylococcus aureus plasmid pT181 (Novick et al. 1989). In its natural form the attenuator is the sequence in the 5' untranslated region of a pT181-encoded mRNA for the plasmid replication protein RepC. Attenuators use RNA structural changes to repress transcription in response to an antisense RNA (Figure 1.4). A number of RNA engineering strategies have utilized the pT181 attenuator as a starting point. Since the attenuator was thought to largely regulate transcription (Brantl & Wagner 2002), initial engineering efforts used a transcriptional fusion following Brantl and Wagner's work that included a portion of the repC coding sequence followed by a stop codon and a separate ribosome binding site for the downstream gene of interest. In this configuration, poor repression was observed which motivated engineering the terminator sequence to increase transcriptional repression from 64% to 85% by the addition of GC pairs (Lucks et al. 2011). This was then used to build a library of orthogonal repressors (Takahashi & Lucks 2013) and the mechanism was reversed to build activators called small transcription activating RNAs (STARs) (Chappell, Takahashi, et al. 2015). These attenuators were also engineered to respond to small molecule and protein

signals by fusing a theophylline aptamer (small molecule) or an MS2 (bacteriophage coat protein) aptamer to the antisense making them theophylline or MS sensitive (Qi et al. 2012). Genetic circuits have been constructed with these orthogonal regulators including logic gates(Chappell, Takahashi, et al. 2015), (Lucks et al. 2011) transcriptional cascades (Lucks et al. 2011), and single input modules (Takahashi et al. 2015). These circuits have been shown to be faster than protein-based networks due to their dependence on the degradation rate of RNA which is typically much faster than protein degradation (Takahashi et al. 2015) and models have been developed to predict their behavior (Hu et al. 2015; Hu et al. 2018).

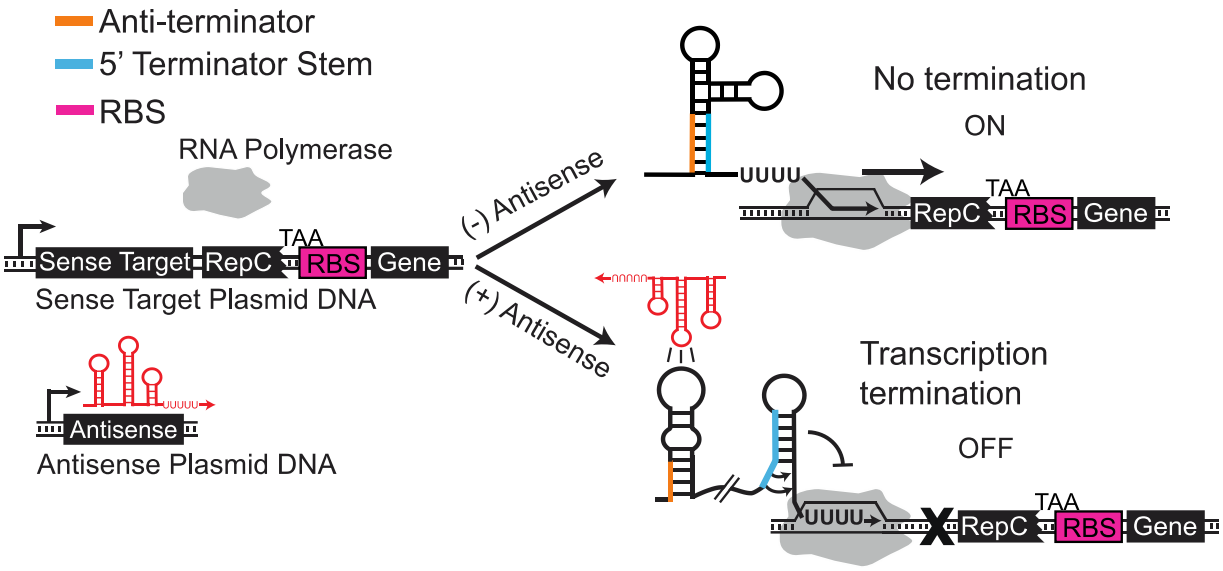


Figure 1.4 Schematic of the pT181 transcriptional repression mechanism. The pT181 attenuator sense target sequence resides in the 5' untranslated region and regulates the expression of a downstream gene. The natural attenuator encoded in plasmid pT181 regulates the expression of the repC gene (Kumar & Novick 1985). Following the attenuator sequence (Brantl & Wagner 2000), 96 nt fragment of the repC gene ending in a stop codon, TAA, is included as a transcriptional fusion before a ribosome binding site (RBS) and the regulated gene of interest. In the absence of antisense RNA (red), the attenuator folds such

that the anti-terminator sequence (orange) sequesters the 5' region of the terminator stem (blue), preventing terminator formation and allowing transcription elongation by RNA polymerase (grey) causing the attenuator to be transcriptionally ON. When antisense RNA is present, its kissing hairpin interaction with the attenuator sequesters the anti-terminator, thus allowing terminator formation which prevents downstream transcription making the attenuator transcriptionally OFF.

1.4 Construction of RNA-based genetic networks and their parts

The work presented in this thesis focuses on the development and optimization of RNA regulators and their subsequent use in novel RNA-based genetic networks.

1.4.1 Chapter 2: Reducing circuit leak using dual transcriptional and translational regulation

Predictable control of gene expression is vital for the accurate construction of genetic networks. However, many RNA regulators suffer from incomplete repression in their OFF state. This circuit leak can propagate through a network and interfere with network function. The work presented in this chapter addresses this by introducing an improvement to the pT181 attenuator system that takes advantage of natural translational control in addition to its well characterized transcriptional control. By adding this second layer of control, we manage to severely decrease regulator leak, increasing repression from 85% to 98% and activation from 10 fold activation to over 900 fold activation. Additionally, we show that orthogonal version of these dual control attenuators can be constructed by engineering

minimal antisense RNAs. Finally, we use the dual control attenuator in an small molecule-activated RNA-only cascade to demonstrate that it successfully reduces leak in a genetic network. Moving forward, we anticipate these regulators will allow the construction of more accurate RNA-based genetic networks.

1.4.1 Chapter 3: Using RNA sequestration to shift a transfer function

Large libraries of genetic parts have now been built and characterized. The ability to tune these parts can facilitate optimization of new genetic networks. Here we show a method of tuning the pT181 attenuator using a sequester that binds to the antisense and blocks its ability to repress gene expression. We start by using a cell free transcription-translation (TXTL) system and an Echo liquid handler to rapidly characterize the sequesters showing that they can efficiently de-repress gene expression. We are also able to shift the pT181 attenuator transfer function by introducing a threshold with sequestration. After demonstrating TXTL functionality, we tested our sequester system in *E. coli* and showed that pT181 sequestration is able to de-repress gene expression as well as manipulate the relationship between in the input and the output of an inducible system. The ability to shift a transfer function could be vital for tuning future genetic networks, particularly changing the threshold of activation for biosensors.

1.4.1 Chapter 4: Generating a pulse of gene expression using regulatory timescale differences

In order to build sophisticated genetic networks with new behaviors parts from different libraries with unknown compatibilities must be combined predictably. In this chapter we

use TXTL-based characterization experiments and an ODE-based mechanistic model to predict the combined dynamics of two different regulator types: CRISPRi and STARs. We rapidly parameterize the ODE-based model using data generated with TXTL, extract parameters using Bayesian inference, and show STAR and CRISPRi based regulation occur on different timescales. We then combine them into a single circuit that produces a pulse of gene expression using this timescale discrepancy. Our results suggest that these regulator types are modular and are ideal for use in sophisticated dynamic gene networks.

1.4.1 Chapter 5: Enhancing cellular communication by decoding signal pulses

As genetic networks become more sophisticated, one of the challenges is to create ways to pass complex messages to these engineered cellular systems. This chapter focuses on expanding cellular communication methods by proposing two networks that can decode pulses of light or chemical signal. The first network uses protein as genetic 'memory' in an RNA circuit due to differing degradation timescales. Our results suggest that the timescale difference is not enough to efficiently remember state especially when factoring in dilution due to cell division. The second circuit 'remembers' state using bistability and is based off of interconnected toggle switches. A stability analysis suggests that this circuit should remember state under some conditions. Preliminary experiments also suggest that the best method for sending signals is using light-based induction due to experimental challenges using chemical inducers.

1.4.1 Chapter 6: Towards RNA-based biological control

Biological noise can cause uncertainty in engineered biological systems. For networks that rely on constant component concentrations, this variability can lead to unexpected behaviors. In industrial systems, controller are used to maintain a constant output in the face of changing environmental conditions. Here we explore two RNA-based biological controller designs. Both controllers rely on RNA sequestration to perform error calculation but one performs the correction using activation while the other uses negative autoregulation. Previous work has suggested RNA sequestration must be much faster than RNA degradation in order for the controller to function. We find activator sequestration to be less efficient than repressor sequestration. Preliminary data indicates the repression based controller is able to perform reference tracking.

1.5 REFERENCES

Anderson, J.C., Voigt, C.A. & Arkin, A.P., 2007. Environmental signal integration by a modular AND gate. Molecular Systems Biology, 3(1), p.133.

Ausländer, S. et al., 2012. Programmable single-cell mammalian biocomputers. Nature, 487(7405), pp.123–127.

Basu, S. et al., 2005. A synthetic multicellular system for programmed pattern formation. Nature, 434(7037), pp.1130–1134.

Beerli, R.R. & Barbas, C.F., 2002. Engineering polydactyl zinc-finger transcription factors. Nature biotechnology, 20(2), pp.135–141.

Brantl, S., 2007. Regulatory mechanisms employed by cis-encoded antisense RNAs. Current opinion in microbiology, 10(2), pp.102–109.

Brantl, S. & Wagner, E.G., 2000. Antisense RNA-mediated transcriptional attenuation: an in vitro study of plasmid pT181. Molecular microbiology, 35(6), pp.1469–1482.

Brantl, S. & Wagner, E.G.H., 2002. An antisense RNA-mediated transcriptional attenuation mechanism functions in Escherichia coli. Journal of bacteriology, 184(10), pp.2740–2747.

Breaker, R.R., 2012. Riboswitches and the RNA world. Cold Spring Harbor perspectives in biology, 4(2), pp.a003566–a003566.

Carrier, T.A. & Keasling, J.D., 1997. Controlling messenger RNA stability in bacteria: strategies for engineering gene expression. Biotechnology progress, 13(6), pp.699–708.

Carrier, T.A. & Keasling, J.D., 1999. Library of synthetic 5' secondary structures to manipulate mRNA stability in Escherichia coli. Biotechnology progress, 15(1), pp.58–64.

Chappell, J. et al., 2017. Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. Nature Communications, 8(1), p.795.

Chappell, J. et al., 2013. The centrality of RNA for engineering gene expression. Biotechnology journal, 8(12), pp.1379–1395.

Chappell, J., Takahashi, M.K. & Lucks, J.B., 2015. Creating small transcription activating RNAs. Nature chemical biology, 11(3), pp.214–220.

Chappell, J., Watters, K.E., et al., 2015. A renaissance in RNA synthetic biology: new mechanisms, applications and tools for the future. Current opinion in chemical biology, 28, pp.47–56.

Elowitz, M.B. & Leibler, S., 2000. A synthetic oscillatory network of transcriptional regulators. Nature, 403(6767), pp.335–338.

Espah Borujeni, A., Channarasappa, A.S. & Salis, H.M., 2014. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. Nucleic acids research, 42(4), pp.2646–2659.

Gardner, T.S., Cantor, C.R. & Collins, J.J., 2000. Construction of a genetic toggle switch in Escherichia coli. Nature, 403(6767), pp.339–342.

Green, A.A. et al., 2017. Complex cellular logic computation using ribocomputing devices. Nature, 548(7665), pp.117–121.

Green, A.A. et al., 2014. Toehold switches: de-novo-designed regulators of gene expression. Cell, 159(4), pp.925–939.

Gupta, A. et al., 2017. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. Nature biotechnology, 35(3), pp.273–279.

Hu, C.Y. et al., 2018. Engineering a Functional small RNA Negative Autoregulation Network with Model-guided Design. ACS synthetic biology, p.acssynbio.7b00440.

Hu, C.Y., Varner, J.D. & Lucks, J.B., 2015. Generating Effective Models and Parameters for RNA Genetic Circuits. ACS synthetic biology, 4(8), pp.914–926.

Isaacs, F.J. et al., 2004. Engineered riboregulators enable post-transcriptional control of gene expression. Nature biotechnology, 22(7), pp.841–847.

Khalil, A.S. & Collins, J.J., 2010. Synthetic biology: applications come of age. Nature reviews. Genetics, 11(5), pp.367–379.

Kittle, J.D. et al., 1989. Insertion sequence IS10 anti-sense pairing initiates by an interaction between the 5' end of the target RNA and a loop in the anti-sense RNA. Journal of Molecular Biology, 210(3), pp.561–572.

Kumar, C.C. & Novick, R.P., 1985. Plasmid pT181 replication is regulated by two countertranscripts. Proceedings of the National Academy of Sciences of the United States of America, 82(3), pp.638–642.

Liu, C.C. et al., 2011. Regulation of transcription by unnatural amino acids. Nature biotechnology, 29(2), pp.164–168.

Lucks, J.B. et al., 2011. Versatile RNA-sensing transcriptional regulators for engineering genetic networks. Proceedings of the National Academy of Sciences of the United States of America, 108(21), pp.8617–8622.

Machens, F. et al., 2017. Synthetic Promoters and Transcription Factors for Heterologous Protein Expression in Saccharomyces cerevisiae. Frontiers in bioengineering and biotechnology, 5, p.63.

Moon, T.S. et al., 2012. Genetic programs constructed from layered logic gates in single cells. Nature, 491(7423), pp.249–253.

Morfeldt, E. et al., 1995. Activation of alpha-toxin translation in Staphylococcus aureus by the trans-encoded antisense RNA, RNAIII. The EMBO journal, 14(18), pp.4569–4577.

Novick, R.P. et al., 1989. pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. Cell, 59(2), pp.395–404.

Olson, E.J. et al., 2014. Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. Nature methods, 11(4), pp.449–455.

Potvin-Trottier, L. et al., 2016. Synchronous long-term oscillations in a synthetic gene circuit. Nature, 538(7626), pp.514–517.

Purnick, P.E.M. & Weiss, R., 2009. The second wave of synthetic biology: from modules to systems. Nature reviews. Molecular cell biology, 10(6), pp.410–422.

Qi, L. et al., 2012. Engineering naturally occurring trans-acting non-coding RNAs to sense molecular signals. Nucleic acids research, 40(12), pp.5775–5786.

Qi, L.S. et al., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152(5), pp.1173–1183.

Stricker, J. et al., 2008. A fast, robust and tunable synthetic gene oscillator. Nature, 456(7221), pp.516–519.

Tabor, J.J. et al., 2009. A synthetic genetic edge detection program. Cell, 137(7), pp.1272–1281.

Takahashi, M.K. & Lucks, J.B., 2013. A modular strategy for engineering orthogonal chimeric RNA transcription regulators. Nucleic acids research, 41(15), pp.7577–7588.

Takahashi, M.K. et al., 2015. Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. ACS synthetic biology, 4(5), pp.503–515.

Tang, J. & Breaker, R.R., 1997. Rational design of allosteric ribozymes. Chemistry & biology, 4(6), pp.453–459.

Watters, K.E. et al., 2016. Cotranscriptional folding of a riboswitch at nucleotide resolution. Nature structural & molecular biology, 23(12), pp.1124–1131.

CHAPTER 2

ACHIEVING LARGE DYNAMIC RANGE CONTROL OF GENE EXPRESSION WITH A COMPACT RNA TRANSCRIPTION-TRANSLATION REGULATOR

2.1 Abstract

RNA transcriptional regulators are emerging as versatile components for genetic network construction. However, these regulators suffer from incomplete repression in their OFF state, making their dynamic range less than that of their protein counterparts. This incomplete repression causes expression leak, which impedes the construction of larger synthetic regulatory networks as leak propagation can interfere with desired network function. To address this, here we demonstrate how naturally derived antisense RNA-mediated transcriptional regulators can be configured to regulate both transcription and translation in a single compact RNA mechanism that functions in *Escherichia coli*. Using in vivo gene expression assays, we show that a combination of transcriptional termination and RBS sequestration increases repression from 85% to 98%, or activation from 10 fold to over 900 fold, in response to cognate antisense RNAs. We also show that orthogonal repressive versions of this mechanism can be created through engineering minimal antisense RNAs. Finally, to demonstrate the utility of this dual control mechanism, we use it to reduce network leak in an RNA-only cascade that activates gene expression as a function of a small molecule input. We anticipate these regulators will find broad use as synthetic biology moves beyond parts engineering to the design and construction of larger and more sophisticated regulatory networks.

This work was originally published in *Nucleic Acids Research* and has been reproduced here. (Westbrook & Lucks 2017)

2.2 Introduction

RNAs are now understood to play broad regulatory roles across the cell (Strobel et al. 2016). As such, synthetic biologists have sought to use these versatile natural systems to create a diverse array of parts that can regulate many aspects of gene expression including transcription (Lucks et al. 2011; Chappell, Takahashi, et al. 2015; Qi et al. 2013), translation (Green et al. 2014; Pardee et al. 2016), and mRNA degradation (Carrier & Keasling 1999; Carothers et al. 2011; Pfleger et al. 2006). Antisense-mediated RNA transcriptional regulators are particularly versatile because they regulate RNA synthesis as a function of an RNA input and thus can be used to create RNA-only genetic networks (Lucks et al. 2011; Takahashi et al. 2015). RNA genetic networks have many potential advantages over protein-based networks including the possibility of leveraging advances in RNA folding algorithms and design rules for part design (Takahashi et al. 2016; Zadeh et al. 2011) and their natural fast dynamics (Takahashi et al. 2015).

Despite these advantages, RNA transcriptional regulators still suffer from low dynamic range, the ratio of maximum (ON) to minimum (OFF) signal, in comparison to protein-based regulators. Low dynamic range can lead to excess signal in the OFF state, causing networks that contain these regulators to be disrupted by low, transient amounts of gene expression signal called network leak. This leak can propagate through the network causing it to function incorrectly, for example by causing a network to express a gene when repression is desired. Previous research has focused on reducing leak to diminish undesired effects, for example by using a recombinase to control gene availability in a multigene network to construct a digital switch biosensor (Lapique & Benenson 2014). However, low dynamic range still remains a significant barrier to using RNA transcriptional repressors in

large genetic networks. While there has been progress in creating RNA translational activators with low leak (Green et al. 2014), there is still room for improvement in RNA translational repressors (RNA IN/OUT 90% (10-fold) repression (Mutalik et al. 2012)), RNA transcriptional repressors (pT181 and variants 85% (6-fold) repression (Lucks et al. 2011)), and RNA transcriptional activators (STAR's 90 fold activation (Chappell, Takahashi, et al. 2015)). Thus an important challenge for RNA engineering is to improve the dynamic range of RNA regulators so that they can be more effective as elements of synthetic genetic networks.

While there has been great progress in improving the dynamic range of RNA regulators by engineering mechanisms that control a single gene expression process (Lee et al. 2016; Qi et al. 2013; Chappell, Takahashi, et al. 2015), only several studies have explored the idea of engineering multiple genetic control processes for tighter regulation (Morra et al. 2016; Horbal & Luzhetskyy 2016; Liu et al. 2012). Specifically, Morra et al. recently combined transcriptional and translational control with two distinct mechanisms - inducible promoters and orthogonal translational riboswitches - to achieve tight control of fluorescent proteins (Morra et al. 2016). Horbal and Luzhetskyy also recently used a similar approach to control pamamycin production in *Streptomyces albus* (Horbal & Luzhetskyy 2016). Using RNA engineering strategies, Liu et al. pursued a different approach by combining RNA-mediated translation regulators with leader-peptide transcriptional attenuators to create a hybrid RNA mechanism that uses sequential control of translation then transcription to achieve large dynamic range repression and activation (Liu et al. 2012). Importantly this study showed that multiple RNA structures can be combined together to regulate several aspects of gene expression.

A notable feature of RNA regulatory mechanisms is that they regulate transcription, translation, and mRNA degradation through the conditional formation of simple hairpin structures at defined positions in mRNAs (Chappell, Watters, et al. 2015). Specifically, intrinsic transcriptional terminators repress transcription by causing the dissolution of the transcription elongation complex (Ray-Soni et al. 2016; Chen et al. 2013), ribosome binding site (RBS)-sequestering hairpins block translation by inhibiting ribosome binding (Brantl 2007; Espah Borujeni et al. 2014), and stability hairpins can prevent the activity of RNases to control mRNA degradation
(Carrier & Keasling 1997; Alifano et al. 1994). The common connection between structure and function exhibited by RNA regulatory mechanisms reveals an intriguing possibility of engineering hairpin structures that can regulate multiple control points within a single mechanism.

We sought to use this approach on the pT181 attenuator from the *Staphylococcus aureus* plasmid pT181 (Novick et al. 1989), which has previously been shown to be useful for engineering a growing number of RNA networks (Lucks et al. 2011; Takahashi et al. 2015; Hu et al. 2015). In its native form, the pT181 attenuator is encoded in the 5' untranslated region of the plasmid replication protein RepC mRNA. Without the cis-encoded antisense RNA repressor, the sense RNA attenuator folds into a structure that allows for transcription of the RepC mRNA. When the antisense RNA is present, its binding to the sense RNA target permits the formation of an intrinsic transcriptional terminator upstream of the RepC coding sequence thereby terminating transcription (Figure 2.1).

A number of RNA engineering strategies have used the pT181 attenuator as a starting point to create RNA genetic networks and gene expression logics. Earlier studies concluded

that the attenuator primarily regulates transcription (Brantl & Wagner 2002), leading initial engineering efforts to use a transcriptional fusion of the attenuator to create basic RNA transcriptional repressors (Lucks et al. 2011). This transcriptional fusion (Brantl & Wagner 2002) included a fragment of the RepC coding sequence followed by a stop codon and a separate ribosome binding site for translation of the downstream gene of interest after the transcriptional decision was made by the attenuator (Figure 2.1A). This configuration initially exhibited only 64% repression, but was improved by strengthening the base of the terminator (Ray-Soni et al. 2016) through the addition of GC pairs to achieve 85% repression (Lucks et al. 2011). Subsequent work used this system to build a library of independently acting, or orthogonal, transcriptional repressors that only repress their cognate targets with minimal cross talk with other variants (Takahashi & Lucks 2013). Orthogonal pairs of regulators are important for networks to function as expected by only controlling target genes as desired without interfering with off target expression. Recently the pT181 mechanism was used to build RNA transcriptional activators (Chappell, Takahashi, et al. 2015), and a variety of genetic networks including logic gates (Chappell, Takahashi, et al. 2015; Lucks et al. 2011), transcriptional cascades (Lucks et al. 2011), and genetic networks that sequentially activate multiple genes (Takahashi et al. 2015).

Intriguingly, early studies on the natural pT181 attenuator mechanism hypothesized that an AGGAG sequence embedded in the 3' half of the terminator hairpin was the ribosome binding site for *repC* (Kumar & Novick 1985). This would suggest that the terminator hairpin of the pT181 attenuator could also function by occluding the RBS to regulate translation as well as transcription. Later, it was determined that the primary mechanism of repression was transcription by comparing transcriptional versus translational reporter gene fusions

(Brantl & Wagner 2002). However the presence of a near canonical RBS sequence in the 3' terminator hairpin, spaced 12 nt from the start codon of *repC* suggests the possibility that the pT181 mechanism may in fact have a more powerful effect on gene expression by simultaneously regulating transcription and translation through the conditional formation of a single compact hairpin in response to interactions with an antisense RNA (Figure 2.1B).

In this work, we show that antisense-mediated repression of gene expression can be improved by utilizing the native RBS and thus the natural dual transcriptional/translational regulation of the pT181 attenuator. When configured as a translational fusion, we show we can increase the percent repression (100%-OFF gene expression level/ON gene expression level) of a fluorescent reporter protein from 85% (+/- 3.4%) to 98% (+/- 0.4%) in *Escherichia coli*. The success of this strategy led us to use it to improve the fold activation (ON gene expression level/OFF gene expression level) of a small transcription activating RNA (STAR) system based on the pT181 hairpin from 10 fold (+/- 3.7) to 923 fold (+/- 213). Our next goal was to create a library of orthogonal dual control repressors that can function independently in the same cell as components of larger genetic networks. To do this, we converted previously published orthogonal pT181 variants that functioned at the transcriptional level (Takahashi & Lucks 2013) into dual control regulators. Interestingly, this library of dual control repressors showed significant cross-talk, indicating that the dual control system breaks orthogonality, likely by increasing the opportunity for non-cognate antisense RNAs to bind and induce translational repression. To mitigate this, we engineered a minimal antisense RNA that reduced crosstalk thereby allowing the repressors to function independently. Finally, to demonstrate that these regulators can be used to fix leak within RNA genetic networks and that orthogonal versions can function in the same cell without

breaking network function, we constructed a repressor cascade using the dual control repressor on the bottom level and found that the dual control cascade exhibited reduced network leak and a higher dynamic range.

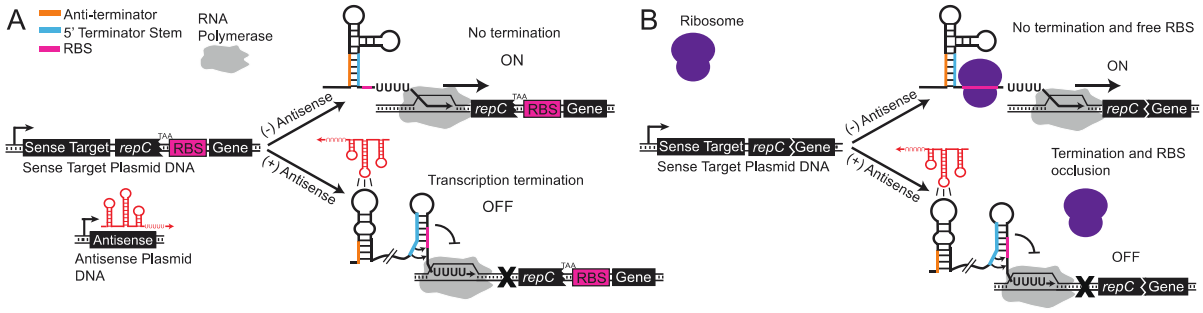


Figure 2.1. Schematic of the transcriptional pT181 repression mechanism (A) and the proposed pT181 dual transcription/translation repression mechanism (B). The pT181 attenuator sense target sequence resides in the 5' untranslated region and regulates the expression of a downstream gene. The natural attenuator encoded in plasmid pT181 regulates the expression of the repC gene (Kumar & Novick 1985). For the transcriptional fusion, a 96 nt fragment of the *repC* gene ending in a stop codon, TAA, is included after the attenuator sequence (Brantl & Wagner 2000) and before a ribosome binding site (RBS) and the regulated gene of interest. For the dual control/translational fusion,12 nt of the *repC* gene is included and is translationally fused to the regulated gene of interest. In the absence of antisense RNA (red), the attenuator folds such that the anti-terminator sequence (orange) sequesters the 5' region of the terminator stem (blue), preventing terminator formation and allowing transcription elongation by RNA polymerase (grey). Thus in the absence of antisense RNA the attenuator is transcriptionally ON. In the dual control/translational fusion, this structure also contains an exposed ribosome binding site

(RBS) for the gene of interest, which allows ribosomes (purple) to bind and translate the mRNA. When antisense RNA is present, its kissing hairpin interaction with the attenuator sequesters the anti-terminator, thus allowing terminator formation, which prevents downstream transcription. Thus in the presence of antisense RNA the attenuator is transcriptionally OFF. The dual control version is both transcriptionally and translationally off in this case due to the added effect of RBS occlusion by the terminator hairpin. Sequences and structures for the dual control attenuator are shown in Supplementary Figure A8.

2.3 Results

2.3.1 Regulating both transcription and translation with a single RNA structure improves dynamic range

We first sought to evaluate the performance of the dual control repressor by configuring it as a translational fusion with a downstream reporter gene (Figure 2.1B). Because the terminator hairpin contains a canonical RBS in its 3' half, we would expect this configuration to regulate both transcription *and* translation of the downstream gene. Specifically, in the presence of antisense RNA, the formation of the terminator hairpin should both repress transcription of the downstream gene, as well as occlude the initiation of translation of any mRNA transcripts that were extended due to imperfect termination efficiency. Thus, we expected the dual control translational fusions to exhibit lower OFF levels than the transcription-only regulators.

In previous work, a translational fusion of the pT181 attenuator to the *lacZ* gene exhibited 62% repression in the presence of an antisense RNA as measured by Miller assays

(Brantl & Wagner 2002). Since the terminator of the pT181 system had been previously engineered (Lucks et al. 2011) to strengthen the terminator stem base in order to increase transcriptional repression (Ray-Soni et al. 2016), we began by assessing the observed antisense-mediated repression of both the natural and engineered terminator using a translational fusion between *repC* and an SFGFP reporter gene (Figure 2.2A,B). To characterize attenuator function, plasmids were constructed such that each attenuator was placed downstream of a constitutive promoter and upstream of the SFGFP coding sequence on a medium copy plasmid. Complementary antisense RNAs were placed on a separate high copy plasmid downstream of the same constitutive promoter (Supplementary Table A1). Each attenuator plasmid was transformed into *E. coli* TG1 cells along with either its cognate antisense or a no-antisense control plasmid (Supplementary Table A2). Individual colonies were picked, grown overnight, sub-cultured into minimal media and grown until logarithmic growth was reached. Fluorescence was measured for each culture using flow cytometry (see materials and methods). Using this experimental design, we observed a 63% (+/- 7.9%) repression in gene expression for the wild-type transcriptional fusion that increased to 98% (+/- 0.4%) when a translational fusion was used (Figure 2.2A). A closer examination of the increase in repression revealed that the translational fusion not only decreased the OFF level of gene expression in the presence of antisense, but also increased the ON level in the absence in antisense.

We performed the same fluorescence experiment described above with the engineered terminator and found an improvement from 85% (+/- 3.4%) repression to 98% (+/- 0.7%) repression (Figure 2.2B). However, in this case the ON level was reduced for the translational fusion, which could be due to the terminator mutation causing increased

spacing between the RBS and the start codon of *repC*. This suggested that the *repC* context of the dual control system could be important. To test this, we fully removed the *repC* sequence and characterized the dual control attenuator. We found that when *repC* is fully removed, the ON level is reduced (Supplementary Figure A4), possibly due to the sequence context change. For this reason we chose to continue with the wild type translational fusion repressor including a 12nt fusion of *repC*.

To further investigate the mechanism of the attenuator, we performed qRT-PCR experiments on a transcriptional and a translational fusion construct (Supplementary Figure A5). The transcriptional attenuator with the engineered terminator showed 80% (+/- 3.4%) repression when measuring SFGFP fluorescence and 78% (+/- 9.1%) repression when measuring SFGFP transcripts with qRT-PCR, indicating that repression comes primarily from transcriptional termination. The dual control attenuator with the wild type terminator showed 97% (+/- 0.7%) repression when measuring SFGFP fluorescence and 84% (+/- 8.1%) repression when measuring SFGFP transcripts with qRT-PCR, indicating that the increased repression is due to the added translational control.

We next designed a construct to compare dual transcription/translation control to transcription-only control using a dual reporter protein operon (Figure 2.2C). In this design, mRFP is translationally fused to the attenuator, while SFGFP is translated from an independent downstream RBS. In this way, we would expect mRFP to be regulated at both the transcriptional and translational levels, while SFGFP would be regulated at just the transcriptional level leading to overall increased repression for mRFP. We transformed cells with the sense target plasmid and the antisense repressor or a blank control plasmid and measured the fluorescence using flow cytometry. As expected, we found that mRFP was

repressed more effectively (92% +/- 1.7%) than SFGFP (71% +/- 5.8%). This result also demonstrated that the dual control repressor can be modularly used to regulate different proteins as well as operons.

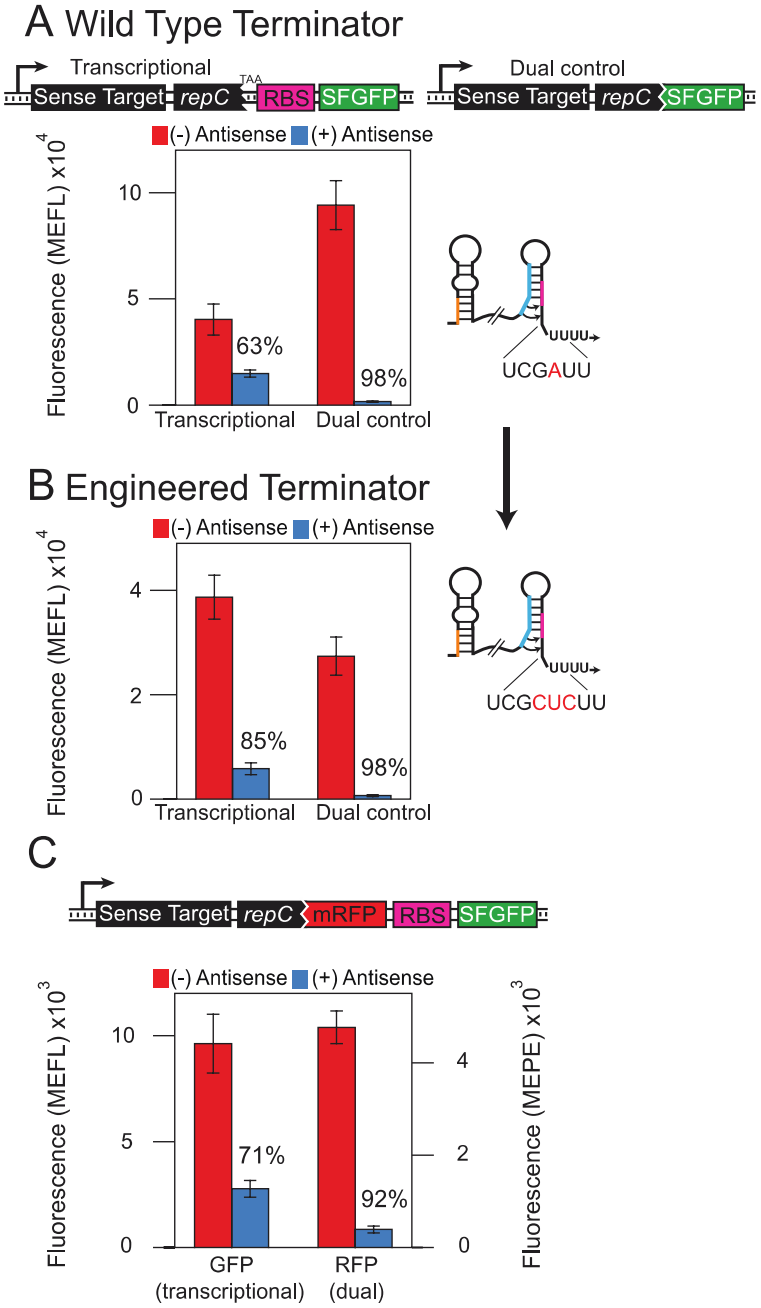


Figure 2.2. Dual transcription/translation control represses gene expression with higher dynamic range than transcription control in vivo. Functional characterization of the (A) wild type (Brantl & Wagner 2002), or (B) engineered (Lucks et al. 2011) attenuator configured to repress either transcription (transcriptional fusion) or dual transcription/translation (translational fusion) of an SFGFP coding sequence. Average fluorescence was collected by flow cytometry as Molecules of Equivalent Fluorescein (MEFL) of *E. coli* TG1 cells transformed with a plasmid expressing the indicated attenuator-SFGFP construct and a plasmid expressing the antisense RNA (+, blue) or a control plasmid lacking the antisense sequence (-, red) (Supplementary Table A2). Percent repression is labelled above each construct tested. In both cases the dual control regulator showed 98% repression (50-fold), though with a higher ON expression level for the wild type attenuator. Error bars represent standard deviations of at least seven biological replicates. Cartoons highlight differences between the wild type and engineered attenuator sequences, which differ by several bases in the 3' half of the terminator hairpins. (C) Testing dual control vs. transcriptional control in a two-colour operon construct. The wild type attenuator sequence was translationally fused to an mRFP coding sequence, which was followed by an RBS-SFGFP sequence. In this way mRFP was under dual transcription/translation control while SFGFP was under only transcription control. The construct was tested as in (A) with mRFP fluorescence collected by flow cytometry as Molecules of Equivalent Phycoerythrin (MEPE). RFP was more strongly repressed at 92% (+/- 1.7%) than GFP at 71% (+/- 5.8%). Averages and standard deviations plotted in (A) and (B) are presented Supplementary Table A3 to allow for comparison within orders of magnitude.

2.3.2 The dual control strategy can be extended to a pT181-based activator to dramatically improve fold activation

We next sought to determine if the dual control strategy could be applied to an RNA-based transcriptional activator mechanism derived from the pT181 system. Small transcription activating RNAs (STARs) were recently engineered to activate, rather than repress, transcription in the presence of designed antisense RNAs (Chappell, Takahashi, et al. 2015). In the STAR mechanism, the sense target region consists of a transcriptional terminator placed upstream of a target gene which blocks transcription elongation to form the OFF state in the absence of a STAR antisense RNA (Supplementary Figure A6). The addition of a STAR antisense RNA, designed to contain an anti-terminator sequence complementary to the 5' half of the terminator stem, prevents terminator formation, allowing transcription to proceed and gene expression to be ON. Early investigations showed that the pT181 attenuation system could be converted into a STAR by using the terminator sequence from pT181 and an appropriately designed STAR antisense RNA (Chappell, Takahashi, et al. 2015). This gave us the opportunity to examine whether a dual control strategy would be effective in the context of gene expression activation.

We constructed a dual control activator by making a translational fusion using one of the pT181 STARs (Supplementary Table A1) (Figure 2.3A). To characterize dual control and transcription-only STAR activator function, each sense target plasmid was transformed into *E. coli* TG1 cells along with either its cognate STAR antisense or a no-antisense control plasmid (Supplementary Table A2). Individual colonies were picked, grown overnight, sub-cultured into minimal media and grown until logarithmic growth was reached. Fluorescence was measured for each culture using flow cytometry (see Materials and Methods). The dual

control strategy improved transcription-only activation from 10 fold (+/- 3.7) to 923 fold (+/- 213) respectively, due to both a higher ON level and a lower OFF level. Notably the OFF level for the dual-control STAR system was remarkably close to the background cellular autofluorescence level (Figure 2.3C).

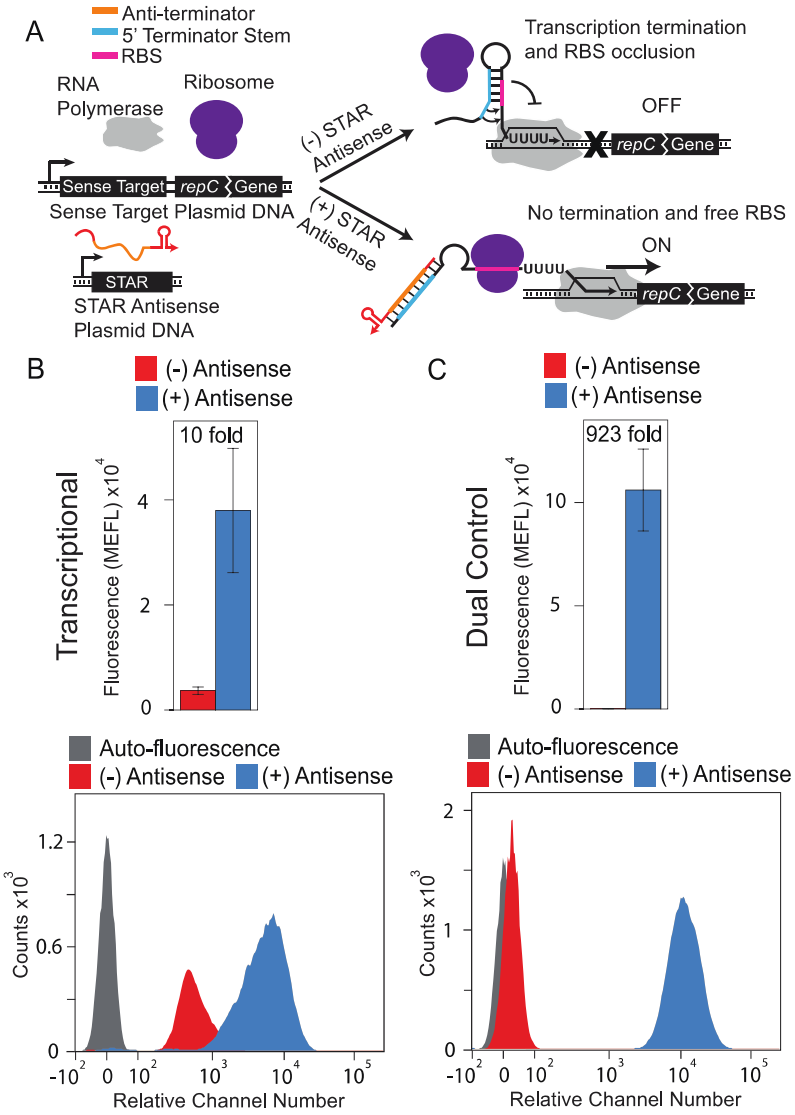


Figure 2.3. Converting a small transcription activating RNA (STAR) mechanism to a dual transcription/translation activator enhances fold activation. (A) Schematic of the proposed

dual transcription/translation activation mechanism. The sense target region consists of the pT181 STAR target region from Chappell et al. (Chappell, Takahashi, et al. 2015) followed by 12 nt of the *repC* gene translationally fused to SFGFP. In the absence of the STAR RNA (red/orange), the terminator forms, preventing downstream transcription by RNA polymerase (grey). This structure also occludes the RBS inside the 3' side of the terminator hairpin, which prevents ribosome binding. Thus in the absence of STAR RNA the mechanism is transcriptionally and translationally OFF. The STAR RNA contains an anti-terminator sequence (orange) complementary to the 5' half of the terminator (blue). When present, the STAR RNA binds to the terminator, preventing terminator formation and allowing transcription elongation. This structure also exposes the RBS, allowing ribosome binding and translation. Thus in the presence of STAR RNA the mechanism is transcriptionally and translationally ON. The original transcriptional mechanism is shown in Supplementary Figure A6. Sequences and structures are shown in Supplementary Figure A11. (B) Functional characterization of a pT181 STAR that controls transcription. Average fluorescence (MEFL) (top) was collected by flow cytometry of *E. coli* TG1 cells transformed with a plasmid expressing the STAR target transcriptionally fused to an SFGFP coding sequence and a plasmid expressing the STAR RNA (+, blue) or a control plasmid lacking the STAR sequence (-, red) (Supplementary Table A2). Error bars represent standard deviations of at least seven biological replicates. The flow cytometry histogram data (bottom) is plotted on a bi-exponential graph (Parks et al. 2006). Auto-fluorescence indicates the observed fluorescence distribution from *E. coli* TG1 cells transformed with plasmids lacking activator-SFGFP fusion or antisense (Supplementary Table A2) (C) Functional characterization of a pT181 STAR that controls both transcription and

translation. Data was collected and plotted as in (B). The dual control strategy increases fold activation from 10 fold (+/- 3.7) to 923 fold (+/- 213) by increasing the ON expression as well as decreasing the OFF expression to near-background auto-fluorescence levels. Averages and standard deviations plotted in (B) and (C) are presented Supplementary Table A4 to allow for comparison within orders of magnitude.

2.3.3 Multiple dual control regulators can be built using pT181 mutants and chimeras

We next sought to determine if the dual control strategy could be applied to additional transcriptional attenuators to improve their dynamic range. Multiple orthogonal, or independently acting, pairs of antisense/attenuators are needed in order to build more sophisticated genetic networks. Since a library of orthogonal pT181 transcriptional regulators has previously been engineered (Takahashi & Lucks 2013), we first sought to apply the dual control strategy to these additional regulators. To create orthogonal antisense/attenuator pairs, the library includes several pT181 specificity changing mutants in the first attenuator hairpin that affect antisense recognition, as well as chimeric fusions of the pT181 mechanism with RNA kissing-hairpin interaction regions taken from translational repressors. However, in order to preserve their overall function, the pT181 mutants and fusions are very similar in sequence, including the pT181 terminator hairpin, allowing us to make translational fusions to test the dual control strategy in these mutant contexts.

Additional dual control repressors were characterized as above and compared to the repression observed in the transcription-only regulatory configuration. Specifically, we tested the transcriptional wild type (WT) repressor, the mutant repressors (Mut 1,2) (Lucks

et al. 2011), and fusion repressors (Fus 3,4) (Takahashi & Lucks 2013) and observed between 63% and 85% repression (Figure 2.4B). We then tested the dual control repressors made from the same attenuators and found that repression increased to between 81% and 98% (Figure 2.4C) averaging to a 15% increase in repression with the wild type pT181 remaining the best dual control repressor. As before, these increases in dynamic range come from both a higher ON level and a lower OFF level (Figure 2.4).

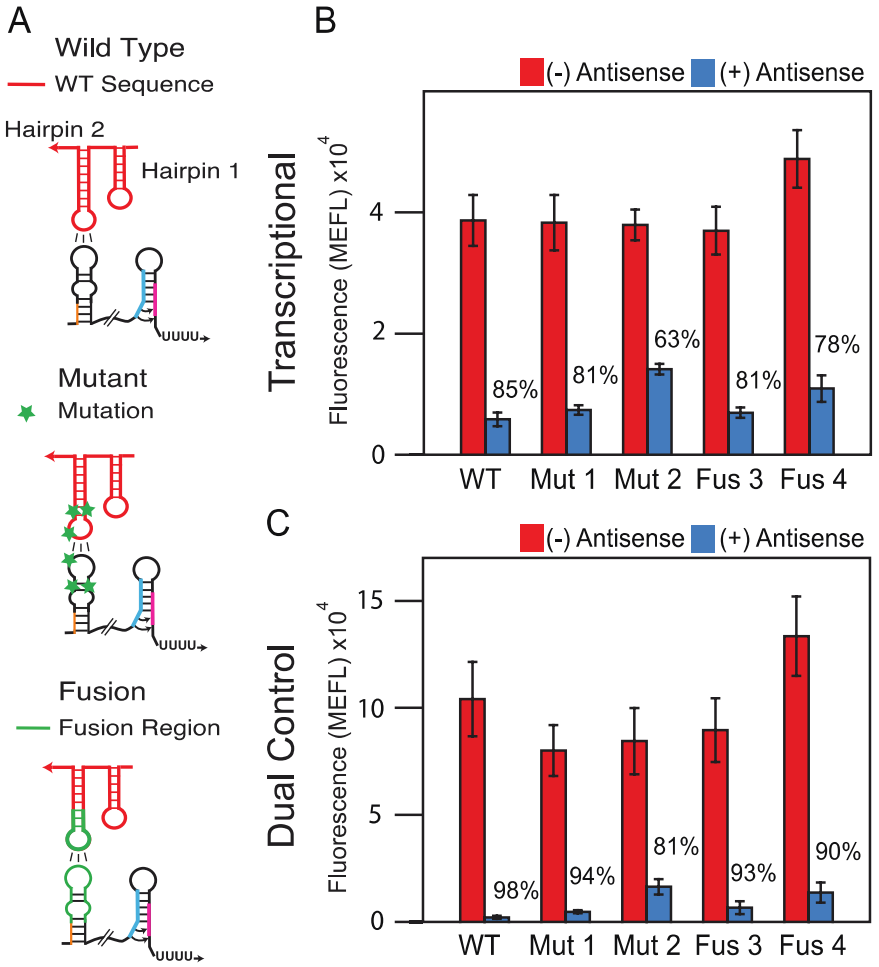


Figure 2.4. The dual transcription/translation control strategy functions across orthogonal pT181 mutants and chimeras. (A) Schematics of the interactions between the dual control

sense target region and the corresponding cognate antisense RNA for wild type, specificity mutants and chimeric fusions engineered to change the specificity of the antisense-attenuator interactions. Sequences and structures are shown in Supplementary Figure A12. (B) Functional characterization of the transcriptional wild type pT181 repressor (WT), two mutants (Mut 1,2) (Lucks et al. 2011), and two chimeric fusions (Fus 3,4) (Takahashi & Lucks 2013). Each repressor contained the wild type terminator region depicted in Figure 2.2A. Functional characterization and data presentation as in Figure 2.2. Error bars represent standard deviations of at least seven biological replicates. (C) As in (B) except with each repressor configured as a dual transcription/translation controller. Using the dual control strategy improves the repression of the transcriptional attenuators. Averages and standard deviations plotted in (B) and (C) are presented Supplementary Table A5 to allow for comparison within orders of magnitude.

2.3.4 Orthogonal dual control repressors can be engineered by reducing the antisense RNA sequence

We next sought to test the orthogonality of the dual control repressors. In addition to requiring multiple dual control regulators to build genetic networks, these regulators must also be orthogonal, or only interact with their cognate target. Previous work showed that the original transcription-only chimeric fusions exhibited limited crosstalk between non-cognate antisense/sense target pairs, making them highly orthogonal (Takahashi & Lucks 2013). To test this for our dual control repressors, we challenged each repressor sense target with all non-cognate antisense RNAs to form an orthogonality matrix (Figure 2.5B, Supplementary Figure A7A).

Despite starting from a set of highly orthogonal transcriptional repressors, we observed significant crosstalk between the dual control regulators. Earlier work on elucidating the mechanism of antisense-mediated translation repression suggested that flanking sequences in the antisense RNA can form extended interactions with the sense target RNAs (Kolb et al. 2000). We thus hypothesized that portions of the antisense RNAs can be interacting with the sense target to repress translation even after the transcriptional regulatory decision has been made. To test this hypothesis, we truncated the antisense RNA sequence to the elements necessary for initial RNA-RNA kissing-hairpin interactions that were shown to be essential for the transcriptional regulatory decision (Takahashi et al. 2016). Specifically, hairpin 2 of the pT181 antisense makes contact with the first hairpin of the sense target region of the attenuator that contains the anti-terminator (Supplementary Figure A8). We hypothesized that we could remove the antisense hairpin 1 and truncate the stem of hairpin 2 to reduce cross-talk between the dual control repressors (Figure 2.5A, Supplementary Figure A9). Using these reduced antisense RNAs, we repeated the orthogonality matrix and observed that crosstalk was reduced for most non-cognate interactions (Figure 2.5C, Supplementary Figure A7B). However, not all crosstalk was reduced. Notably sense/antisense pairs that began with low crosstalk values, displayed increased crosstalk using the truncated antisense. For example the pair consisting of fusion 4 antisense targeting the fusion 3 sense target rose significantly from 7% to 27% to become the highest crosstalk. However, generally those non-cognate pairs that started at higher crosstalk were more significantly reduced (reduced 21% on average for those above 20% crosstalk) compared to those that started at lower crosstalk (raised 0.5% on average for those below 20% crosstalk).

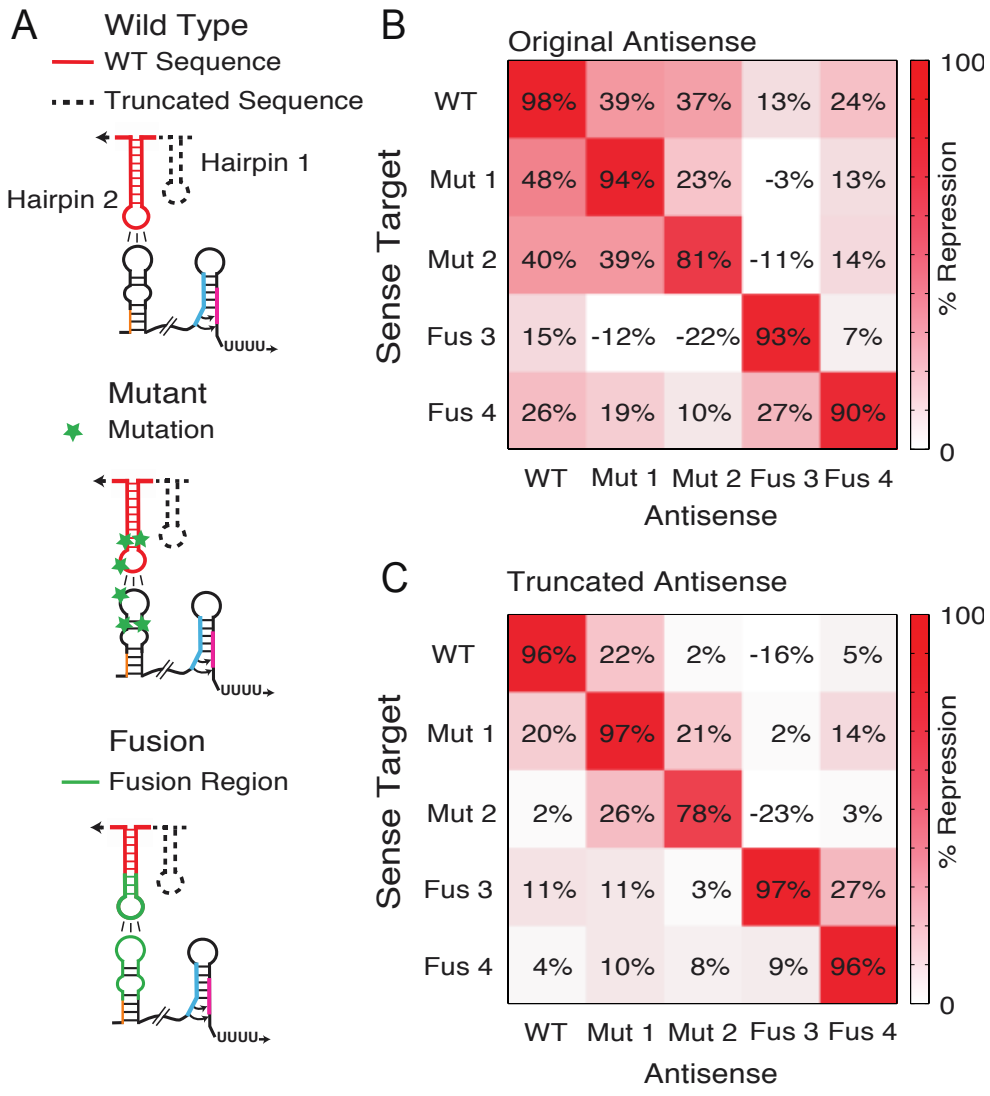


Figure 2.5. Truncated antisense RNA improves orthogonality between dual transcription/translation RNA repressors. (A) Schematics of the interactions between the dual control sense target region and the corresponding cognate antisense RNA for wild type, specificity mutants and chimeric fusions. Dashed lines show portions of the antisense RNA structure that were truncated to reduce cross talk between pairs of dual transcription/translation control RNA repressors. Hairpin 1 and unnecessary regions (4 nt) at the base of hairpin 2 of the antisense were deleted. Sequences and structures for the wild

type pT181 antisense are show in Supplementary Figure A9. (B) An orthgonality matrix showing percent repression observed when sense targets were co-expressed with different full-length antisense RNAs. Each element of the matrix represents the percent repression observed from the indicated antisense/sense target plasmid combination compared to a no-antisense/sense target plasmid condition using functional characterization experiments as in Figure 2.2. (C) As in (B) with truncated forms of the antisense RNAs depicted in (A), showing reduction in repression when non-cognate truncated antisense is present (off diagonal elements). Barplots depicting the data in (B) and (C) are shown in Supplementary Figure A7. Standard deviations for the data in (B) and (C) are shown in Supplementary Tables A6 and A7.

2.3.5 The dual control repressor mitigates network leak in an RNA repressor cascade

Finally we sought to test the dual control regulators in an RNA-only network context that would demonstrate how reduced leak improves network performance and confirm that orthogonal versions can function correctly in the same cell. RNA repressor cascades were the first RNA-only network built (Lucks et al. 2011) and have been used to highlight the fast speed of RNA genetic networks (Takahashi et al. 2015). The repressor cascade also acts as a modular unit that can be built upon to create more sophisticated networks such as one that controls the timing of a sequence of genes in response to a single input (Takahashi et al. 2015; Alon 2007). However, past attempts at characterizing repressor cascades have revealed that the network leaks due to insufficient repression of the individual regulators resulting in un-desired gene expression in conditions where the overall network is designed to be OFF. We therefore sought to fix the leak of an RNA repressor cascade using the dual

control repressor. To test this, we built an RNA repressor cascade that activates the expression of SFGFP in response to theophylline (Figure 2.6A). Theophylline was chosen because a theophylline activated antisense allows us to build an RNA-only network that can directly respond to this small molecule (Qi et al. 2012). However, we also verified that other inducible promoter systems can be used to induce antisense expression and tune repression (Supplementary Figure A10). The cascade consists of three plasmids each expressing one level of the network. Without theophylline present, antisense repressor RNA 2 represses sense target RNA 2 and SFGFP expression. When theophylline is added, it activates antisense repressor RNA 1, which is normally non-functional in the absence of theophylline due to a designed interaction between the antisense RNA hairpin and a fused theophylline aptamer (Qi et al. 2012). In this way, theophylline binding allows antisense repressor RNA 1 to repress antisense repressor RNA 2, allowing SGFP to be expressed. Overall, when theophylline is added to the cell culture, an RNA signal induces SFGFP expression.

To compare RNA cascades that use either transcription-only or dual control SGFP expression, we performed time course experiments on *E. coli* cultures that contained the cascade plasmids with either the transcriptional or dual control repressor cascade plasmids for the bottom level of the cascade. After incubating overnight in LB media, the cultures were diluted into M9 supplemented media and incubated for four hours. The cultures were then diluted again into fresh M9 media to a consistent OD and incubated for four more hours. From here, we sampled cultures every 30 minutes to measure SFGFP fluorescence and culture OD over time. Theophylline was added to some cell cultures at the beginning of sampling to measure the cascade response (Figure 2.6B). This experiment was repeated on three separate days, with the first day shown in Figure 2.6B and the other two shown in

Supplementary Figure A2. In addition we performed a version of this experiment from glycerol stocks which showed similar results (Supplementary Figure A3). As expected, when theophylline was introduced to both the transcriptional and dual control cascades, we observed SFGFP activation that continued throughout the rest of the time course. However, the transcriptional version of the network displayed significant leak (Figure 2.6B, green curves) in comparison to the dual control network, which displayed a lower baseline expression (Figure 2.6B, blue curves) and thus a greater dynamic range. The leak in the transcriptional version of the network is a direct result of the leaky transcriptional repressor – even when in the OFF state, terminator read through can lead to translation of downstream transcripts since their RBS is not masked within a secondary structure. In the dual control case however, the masking of the RBS within the terminator structure prevents this aberrant translation leading to reduced OFF states throughout. This result demonstrated that dual control repressors can not only be used in an RNA genetic network, but that their use reduced overall leak through the network to improve its desired function.

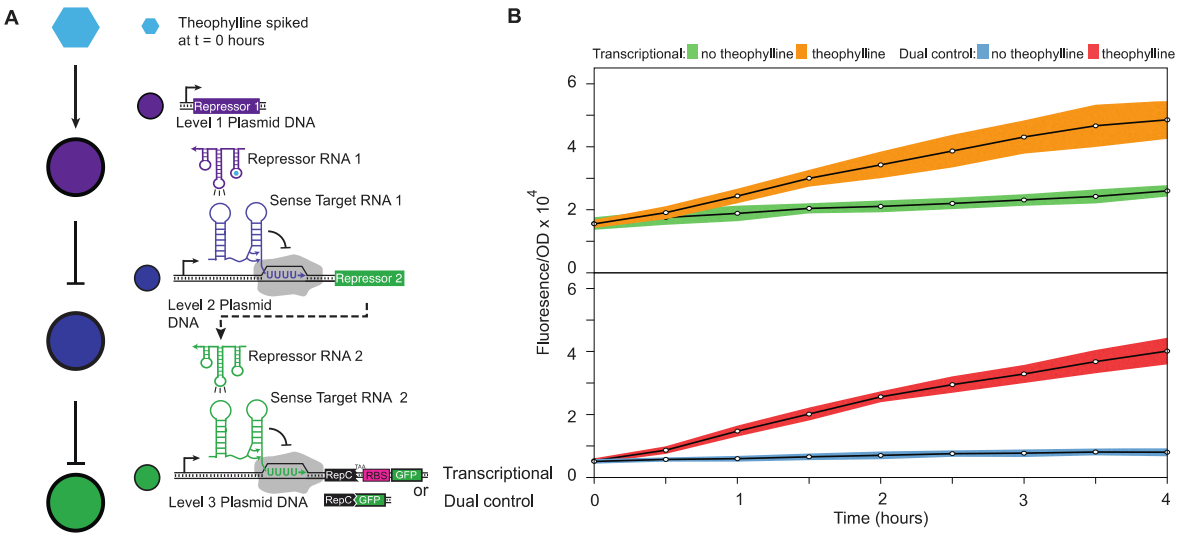


Figure 2.6. The dual transcription/translation control strategy mitigates leak in an RNA

repressor cascade. (A) Schematic of the theophylline activated RNA repressor cascade. The level three SFGFP gene expression is controlled by sense target region 2, which is repressed by repressor RNA 2. Repressor RNA 2 is in turn controlled by the upstream sense target region 1, which is repressed by repressor RNA 1. Repressor RNA 1 is a fusion with a theophylline aptamer (Qi et al. 2012) that is active only with theophylline bound. Without theophylline, repressor RNA 1 is inactive causing overall repression of SFGFP (OFF). When theophylline is added to the cell culture media, the repressor RNA 1 represses transcription of repressor RNA 2, leading to SFGFP expression (ON). The level three attenuator was configured to regulate SFGFP either transcriptionally, or using the dual transcription/translational control mechanism. (B) Functional time course characterization of the transcriptional and dual control repressor cascades. Three plasmids each encoding one of the network levels were co-transformed into *E. coli* TG1 cells, grown overnight and sub-cultured into fresh M9 minimal media for four hours before starting the time-course with a fresh sub-culture (see Methods). After four hours of growth in M9, theophylline (2mM) was added to the media causing SFGFP to be expressed (orange for transcriptional and red for dual control). Time points were sampled every 30 minutes for four hours. Bulk fluorescence and OD600 were measured using a plate reader. The no theophylline condition is shown in green for the transcriptional cascade and blue for dual control. The dual control regulator reduces the overall background fluorescent level while maintaining a similar ON level and thus improves dynamic range. The data shown here are from three individual transformants on a single day. Data for the three independent experiments performed on separate days are shown together in Supplementary Figure A2. The colored region indicates the standard deviation from three biological replicates.

2.4 Discussion

In this work, we have demonstrated the utility of an RNA structure that regulates both transcription and translation in a single, compact mechanism by showing that it improves dynamic range of antisense RNA-mediated control of gene expression and reduces leak when used in RNA genetic networks. Specifically, translational fusions between the pT181 attenuator and downstream reporter genes allowed the transcription of these genes to be regulated by the pT181 terminator hairpin and the translation of these genes by the *repC* RBS sequence encoded in the 3' half of the same hairpin. In this way, the formation of the OFF structure in the presence of a cognate antisense RNA allows gene expression to be repressed at two levels, and thereby improves repression from 85% (+/- 3.4%) for the transcriptional-only case to 98% (+/- 0.4%) in the dual control case. In addition to decreasing OFF levels in the presence of antisense RNA, this configuration increased the ON level in the absence of antisense RNA.

To further investigate the mechanistic details of the dual control repressors we performed qRT-PCR experiments (Supplementary Figure A5). The transcriptional attenuator displayed similar SFGFP fluorescence repression (80%) and SFGFP transcript repression (78%) measured by qRT-PCR, while the dual control attenuator revealed improved SFGFP fluorescence repression (97%) over SFGFP transcript repression (84%). This indicates that some of the improvement is due to the added translational effect of RBS occlusion when the antisense RNA is present. The dual control improvement in repression comes from both an increased ON level in the absence of antisense RNA and a decreased OFF level when antisense is present as compared to the transcriptional attenuator. The increased

ON level could be due to increased RBS exposure due to RNA structural context around the RBS and the start codon with the RNA is in the anti-terminated structure (Supplementary Figure A8). A similar structural effect has recently been seen in the B. cereus *crcB* fluoride riboswitch in which the anti-terminated form shows increased RBS exposure (Watters et al. 2016). In addition, the translational fusion allowed an optimal distance between the RBS and the start codon (Supplementary Figure A8) allowing for more efficient access to the ribosome and greater frequency of translation initiation. We also observed a reduction in the mRNA OFF level in the dual control scenario as compared to the transcriptional attenuator when measuring SFGFP transcripts using qRT-PCR. This effect could be a result of a decrease in translation in the tightly regulated OFF state, which could reduce ribosome protection of the mRNA to allow more efficient mRNA degradation.

Interestingly, our results are different from those observed by previous studies of the pT181 attenuator (Brantl & Wagner 2002). Through comparing transcriptional vs. translational fusions of the pT181 attenuator to the LacZ reporter gene, this study observed 62% repression for the translational fusion and 50% for the transcriptional fusion. The lack of significantly different results and the presence of an intrinsic terminator sequence indicated that the attenuator functioned primarily through transcriptional repression. It is possible that our system shows a more significant difference because of the increased sensitivity afforded by our use of SFGFP expression. Nevertheless, our findings strongly suggest that the natural pT181 attenuator system likely regulates at both the transcriptional and translational levels.

Overall the dual control mechanism significantly improved the dynamic range of RNA regulators over RNA transcriptional repressors and is better than the ~90% repression seen

for the best RNA translational repressors (Mutalik et al. 2012) and the 85% repression seen for the best transcriptional RNA repressors (Lucks et al. 2011). In addition to the pT181 dual control repressor, we also engineered a pT181 STAR activator and increased its activation in response to STAR antisense RNA from 10 fold (+/- 3.7) to 923 fold (+/- 213). This improves upon the previously published fold activation of transcriptional STAR regulators (90 fold (Chappell, Takahashi, et al. 2015)) and translational toehold regulators (~400 fold (Green et al. 2014)). We also showed that this strategy could be expanded to additional pT181 mutant and fusion repressors, increasing the repression of several orthogonal regulators with this unique combination of transcriptional and translation control. Overall this is a significant increase in the number and capability of regulatory tools available for constructing genetic networks with tighter control, which Is particularly useful for situations in which an RNA part with reduced leak is desired.

In order to build robust genetic networks in which the parts act independently and predictably the parts must be orthogonal or act independently of other regulators in the system. However, the initial dual control riboregulators exhibited significant crosstalk. We hypothesized that this was due to additional interactions between the antisense RNAs and the sense target RNAs that caused translation to be repressed even after the transcriptional decision had been made. For a transcriptional decision to be made, the antisense RNAs must interact cotranscriptionally. However, the antisense can still bind the dual control sense target after transcription and affect RBS availability. This would indicate that the modifications between mutants, fusions, and the original pT181 are enough to inhibit crosstalk during transcription but the increased time for antisense-sense target interactions before translation initiation allows shared sequences in non-cognate pairs more opportunity

to interact. We therefore decided to reduce the redundant pT181 sequence to reduce the affinity of non-cognate antisense RNAs for sense target regions. While not all non-cognate antisense and sense target pairs were improved by the truncations, on average the worst performing pairs improved by 20% repression. By truncating redundant pT181 sequence we greatly improved orthogonality, making it possible to use these dual control repressors in RNA networks.

Finally, we used dual control RNA repressors to address a current problem with RNA only networks, which is leak that results from parts that do not allow complete repression of their targets. Specifically, we used the dual control repressor in a repressor cascade and found that it reduced network leak and background fluorescence. Additionally, the repressor cascade demonstrates that the orthogonal attenuators can function together in the same cell.

This work demonstrates a novel RNA motif that regulates multiple aspects of gene expression in a single compact mechanism, and that displays a dynamic range of gene regulation comparable to protein-based mechanisms. As such, this is another example of how RNAs may be optimized to function as well as proteins. We anticipate that as synthetic biology moves beyond the creation of regulator parts libraries and into building more sophisticated networks, RNA regulatory mechanisms such as dual control repressors will find increased use in designing RNA genetic networks with predictable function.

2.5 Materials and Methods

Plasmid construction

Key sequences can be found in Supplementary Table A1. All the plasmids used in this study can be found in Supplementary Table A2 and plasmid diagrams in Supplementary Figure A1.

The pT181 repressor and antisense plasmids, the pT181 mutant repressor and antisense plasmids, and the no-antisense control plasmid were constructs pAPA1272, pAPA1256, pAPA1273, pAPA1257, and pAPA1260, respectively, from Lucks et al. (Lucks et al. 2011). The top level of the cascade was the theophylline pT181 mutant antisense plasmid, construct pAPA1306, from Qi et al. (Qi et al. 2012). The middle level of the cascade was modified from construct pAPA1347 from Lucks et al. (Lucks et al. 2011) using Golden Gate assembly (Engler et al. 2008). The bottom level of the transcriptional cascade was construct pJBL1855 from Takahashi et al. (Takahashi et al. 2015) and the bottom level of the dual control cascade was modified from this construct using Golden Gate assembly. The antisense and repressor plasmids were constructed using inverse PCR (iPCR).

Strains, growth medium, and In Vivo end point gene expression.

All experiments were performed in *E. coli* strain TG1. Experiments were performed for at least seven biological replicates, (independent transformants of an isogenic strain unless otherwise noted) collected over three separate days. Plasmid combinations were transformed into chemically competent *E. coli* TG1 cells, plated on Difco LB+Agar plates containing 100 μg/mL carbenicillin and 34 μg/mL chloramphenicol and incubated overnight at 37 °C. Plates were taken out of the incubator and left at room temperature for approximately 9 h. Three colonies were picked and used to inoculate 300 μL of LB containing carbenicillin and chloramphenicol at the concentrations above in a 2 mL 96-well block (Costar 3960), and grown approximately 17 h overnight at 37 °C at 1,000 rpm in a Labnet Vortemp 56 benchtop shaker. Six microliters of each overnight culture was then added to separate wells on a new block containing 294 μL (1:50 dilution) of supplemented M9

minimal media (1xM9 minimal salts, 1 mM thiamine hydrochloride, 0.4% glycerol, 0.2% casamino acids, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$) containing the selective antibiotics and grown for 4 h at the same conditions as the overnight culture. Cultures (6-12 μL) were then transferred into a FACS round-bottom 96 well plate with 244 μL of PBS containing 2mg/mL Kanamycin to stop translation. The plate was then read on a BD LSR II using the high throughput setting with the high throughput sampler (HTS). The samples for Figure 2.2C were transferred to Falcon 5 ml polystyrene round-bottom tubes and analyzed on a BD Aria Fusion. Growth information for glycerol stocks of an individual transformant for the cascade experiment is described in Supplementary Note A1.

Flow cytometry data collection.

Data for the following parameters were collected on the BD LSR II: forward scatter (FSC), side scatter (SSC), and SFGFP (Pédelacq et al. 2006) fluorescence (488 nm excitation, 515-545 nm emission). Three to ten uL of each sample was measured in high throughput mode. Each sample was required to have at least 5,000 counts, but most had 10,000 to 50,000. Counts were gated in FSC vs. SSC by choosing a window surrounding the largest cluster of cells. SFGFP fluorescence values were recorded in relative channel number (1-262,144 corresponding to 18-bit data) and the geometric mean over the gated data was calculated for each sample. Data for Figure 2.2C was collected on a BD Aria Fusion for the following parameters: forward scatter (FSC), side scatter (SSC), SFGFP fluorescence (488 nm excitation, 530 nm emission), and mRFP Fluorescence (561 nm excitation, 582 nm emission). SFGFP and mRFP fluorescence values were recorded in relative channel number (1-262,144 corresponding to 18-bit data) and the geometric mean over the gated data was calculated

for each sample. Compensation was calculated automatically by the BD Aria FACSDiva software using the compensation setup feature.

Flow cytometry data analysis.

Data analysis and FACS calibration was performed according to the supplementary info of Lucks et al. (Lucks et al. 2011). Spherotech 8-Peak Rainbow Calibration Beads (Spherotech cat. no 559123) were used to obtain a calibration curve to convert fluorescence intensity (geometric mean, relative channel number) into Molecules of Equivalent of Fluorescein (MEFL) units for SFGFP fluorescence or Molecules of Equivalent Phycoerythrin (MEPE) for RFP fluorescence. For each experiment, data for a set of control cultures was also collected which consisted of *E. coli* TG1 cells that do not produce SFGFP (transformed with control plasmids JBL001 and JBL002). The mean MEFL or MEPE value of TG1 cells without SFGFP or mRFP expression, respectively was subtracted from each colony's MEFL or MEPE value. Mean MEFL or MEPE values were calculated over replicates and error bars represent the standard deviation. For repressors, the OFF level is the MEFL or MEPE of cells containing the sense plasmid and the antisense plasmid and the ON level is the MEFL or MEPE of cells containing the sense plasmid and a no-antisense control plasmid. The percent repression for each antisense RNA/attenuator plasmid combination was calculated by subtracting the OFF level divided by the ON level from 1 (1-OFF/ON). For activators the ON level is the MEFL or MEPE of cells containing the sense plasmid and the antisense plasmid and the OFF level is the MEFL or MEPE of cells containing the sense plasmid and a no-antisense control plasmid. The fold activation was calculated by dividing the ON level by the OFF level (ON/OFF).

In Vivo *bulk fluorescence time course experiments*.

Strain, transformation, and media were all the same as for end point experiments described above, except 25ug/mL of kanamycin was used in addition to the other selective antibiotics because the cascade is encoded by three plasmids. Transformation plates containing *E. coli* TG1 cells transformed with three cascade plasmids (Supplementary Table A2) were taken out of the incubator and left at room temperature for approximately 3 h. Three colonies were picked and used to inoculate 300 μL of LB containing selective antibiotics in a 2 mL 96-well block (Costar 3960), and grown approximately 17 h overnight at the same conditions as described for an end point experiment. Twenty microliters of each overnight culture was then added to separate wells on a new block containing 980 μL (1:50 dilution) of supplemented M9 minimal media (as mentioned above) containing the selective antibiotics and grown for 4 h at the same conditions as the overnight culture. The optical density (OD, 600 nm) was then measured by transferring 50 μL of culture from the block into a 96-well plate (Costar 3631) containing 50 μL of phosphate buffered saline (PBS) and measuring using a Biotek Synergy H1m plate reader. The cultures were diluted into 1ml of fresh M9 minimal media to an optical density of 0.015 and grown for four hours. Then theophylline was added to the theophylline condition to a final concentration of 2mM. Every 30 min for the next 4 h, 50 μL from each of the fresh cultures was removed from the 96-well block and transferred to a 96-well plate (Costar 3631) containing 50 μL of phosphate buffered saline (PBS). SFGFP fluorescence (FL, 485 nm excitation, 520 nm emission) and optical density (OD, 600 nm) were then measured at each time point using a Biotek Synergy H1m plate reader.

Bulk fluorescence data analysis.

On each 96-well block, there were two sets of controls; a media blank (M9 alone) and *E. coli* TG1 cells that do not produce SFGFP (transformed with control plasmids JBL001, JBL002, and JBL1856). The block contained three replicates of each control. OD and FL values for each colony at each time point were first corrected by subtracting the corresponding values of the media blank at that same time point. The ratio of FL to OD (FL/OD) was then calculated for each well (grown from a single colony), and the mean FL/OD of TG1 cells without SFGFP at the same time point was subtracted from each colony's FL/OD value to correct for cellular autofluorescence. Experiments were performed for nine biological replicates collected over three separate days. One day is shown in Figure 2.6 while all three days are shown together in Supplementary Figure A2. Data for glycerol stocks of an individual transformant is shown in Supplementary Figure A3 and data collection is described in Supplementary Note A1.

2.6 Acknowledgements

We acknowledge Cameron Glasscock (Julius Lucks Laboratory, Cornell University) for his help in reviewing and checking data processing, Eric Strobel (Julius Lucks Laboratory, Northwestern University) for insightful discussions of the pT181 mechanism, and Sarai Meyer (Julius Lucks Laboratory, Cornell University) for her aid in the revision process in addition to reviewing and checking data processing.

2.7 Funding

This work was supported by the Office of Naval Research Young Investigators Program Award (ONR YIP) [N00014-13- 1-0531 to J.B.L.] an NSF CAREER Award [1452441 to J.B.L], and Searle Funds at The Chicago Community Trust [to J.B.L].

2.8 References

Alifano, P., Bruni, C.B. & Carlomagno, M.S., 1994. Control of mRNA processing and decay in prokaryotes. Genetica, 94(2-3), pp.157–172.

Alon, U., 2007. Network motifs: theory and experimental approaches. Nature reviews. Genetics, 8(6), pp.450–461.

Brantl, S., 2007. Regulatory mechanisms employed by cis-encoded antisense RNAs. Current opinion in microbiology, 10(2), pp.102–109.

Brantl, S. & Wagner, E.G., 2000. Antisense RNA-mediated transcriptional attenuation: an in vitro study of plasmid pT181. Molecular microbiology, 35(6), pp.1469–1482.

Brantl, S. & Wagner, E.G.H., 2002. An antisense RNA-mediated transcriptional attenuation mechanism functions in Escherichia coli. Journal of bacteriology, 184(10), pp.2740–2747.

Carothers, J.M. et al., 2011. Model-driven engineering of RNA devices to quantitatively program gene expression. Science (New York, N.Y.), 334(6063), pp.1716–1719.

Carrier, T.A. & Keasling, J.D., 1997. Controlling messenger RNA stability in bacteria: strategies for engineering gene expression. Biotechnology progress, 13(6), pp.699–708.

Carrier, T.A. & Keasling, J.D., 1999. Library of synthetic 5' secondary structures to manipulate mRNA stability in Escherichia coli. Biotechnology progress, 15(1), pp.58–64.

Chappell, J., Takahashi, M.K. & Lucks, J.B., 2015. Creating small transcription activating RNAs. Nature chemical biology, 11(3), pp.214–220.

Chappell, J., Watters, K.E., et al., 2015. A renaissance in RNA synthetic biology: new mechanisms, applications and tools for the future. Current opinion in chemical biology, 28, pp.47–56.

Chen, Y.-J. et al., 2013. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. Nature methods, 10(7), pp.659–664.

Engler, C., Kandzia, R. & Marillonnet, S., 2008. A one pot, one step, precision cloning method with high throughput capability. H. A. El-Shemy, ed. PLoS ONE, 3(11), p.e3647.

Espah Borujeni, A., Channarasappa, A.S. & Salis, H.M., 2014. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. Nucleic acids research, 42(4), pp.2646–2659.

Green, A.A. et al., 2014. Toehold switches: de-novo-designed regulators of gene expression. Cell, 159(4), pp.925–939.

Horbal, L. & Luzhetskyy, A., 2016. Dual control system - A novel scaffolding architecture of an inducible regulatory device for the precise regulation of gene expression. Metabolic engineering, 37, pp.11–23.

Hu, C.Y., Varner, J.D. & Lucks, J.B., 2015. Generating Effective Models and Parameters for RNA Genetic Circuits. ACS synthetic biology, 4(8), pp.914–926.

Kolb, F.A. et al., 2000. Progression of a loop-loop complex to a four-way junction is crucial for the activity of a regulatory antisense RNA. The EMBO journal, 19(21), pp.5905–5915.

Kumar, C.C. & Novick, R.P., 1985. Plasmid pT181 replication is regulated by two countertranscripts. Proceedings of the National Academy of Sciences of the United States of America, 82(3), pp.638–642.

Lapique, N. & Benenson, Y., 2014. Digital switching in a biosensor circuit via programmable timing of gene availability. Nature chemical biology, 10(12), pp.1020–1027.

Lee, Y.J. et al., 2016. Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. Nucleic acids research, 44(5), pp.2462–2473.

Liu, C.C. et al., 2012. An adaptor from translational to transcriptional control enables predictable assembly of complex regulation. Nature methods, 9(11), pp.1088–1094.

Lucks, J.B. et al., 2011. Versatile RNA-sensing transcriptional regulators for engineering genetic networks. Proceedings of the National Academy of Sciences of the United States of America, 108(21), pp.8617–8622.

Morra, R. et al., 2016. Dual transcriptional-translational cascade permits cellular level tuneable expression control. Nucleic acids research, 44(3), pp.e21–e21.

Mutalik, V.K. et al., 2012. Rationally designed families of orthogonal RNA regulators of translation. Nature chemical biology, 8(5), pp.447–454.

Novick, R.P. et al., 1989. pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. Cell, 59(2), pp.395–404.

Pardee, K. et al., 2016. Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. Cell, 165(5), pp.1255–1266.

Parks, D.R., Roederer, M. & Moore, W.A., 2006. A new “Logicle” display method avoids deceptive effects of logarithmic scaling for low signals and compensated data. Cytometry. Part A : the journal of the International Society for Analytical Cytology, 69(6), pp.541–551.

Pédelacq, J.-D. et al., 2006. Engineering and characterization of a superfolder green fluorescent protein. Nature biotechnology, 24(1), pp.79–88.

Pfleger, B.F. et al., 2006. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nature biotechnology, 24(8), pp.1027–1032.

Qi, L. et al., 2012. Engineering naturally occurring trans-acting non-coding RNAs to sense molecular signals. Nucleic acids research, 40(12), pp.5775–5786.

Qi, L.S. et al., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152(5), pp.1173–1183.

Ray-Soni, A., Bellecourt, M.J. & Landick, R., 2016. Mechanisms of Bacterial Transcription Termination: All Good Things Must End. Annual review of biochemistry, 85(1), pp.319–347.

Strobel, E.J. et al., 2016. RNA systems biology: uniting functional discoveries and structural tools to understand global roles of RNAs. Current opinion in biotechnology, 39, pp.182–191.

Takahashi, M.K. & Lucks, J.B., 2013. A modular strategy for engineering orthogonal chimeric RNA transcription regulators. Nucleic acids research, 41(15), pp.7577–7588.

Takahashi, M.K. et al., 2015. Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. ACS synthetic biology, 4(5), pp.503–515.

Takahashi, M.K. et al., 2016. Using in-cell SHAPE-Seq and simulations to probe structure-function design principles of RNA transcriptional regulators. RNA (New York, N.Y.), 22(6), pp.920–933.

Watters, K.E. et al., 2016. Cotranscriptional folding of a riboswitch at nucleotide resolution. Nature structural & molecular biology, 23(12), pp.1124–1131.

Westbrook, A.M. & Lucks, J.B., 2017. Achieving large dynamic range control of gene expression with a compact RNA transcription-translation regulator. Nucleic acids research, 45(9), pp.5614–5624.

Zadeh, J.N. et al., 2011. NUPACK: Analysis and design of nucleic acid systems. Journal of computational chemistry, 32(1), pp.170–173.

CHAPTER 3

TUNING THE TRANSFER FUNCTION OF AN RNA REPRESSOR USING SEQUESTRATION

3.1 Abstract

RNA regulation is increasingly being used to build complex biological networks, with large libraries of RNA parts of diverse functions reported and characterized. To manipulate and optimize networks constructed from RNA regulators, synthetic biologists must be able to tune the functionality of individual RNA parts. Here we show a method of tuning RNA-mediated repression of gene expression using a sequestration strategy by designing an RNA molecule that binds to and sequesters an antisense RNA repressor. We demonstrate how RNA repressors can be sequestered in both cell free transcription-translation (TXTL) reactions and in *Escherichia coli* to de-repress gene expression. Additionally, we show that multiple sequesters can be made orthogonal, allowing their use in complex biological networks. Finally, we show that the RNA repressor transfer function can be shifted using sequestration in both TXTL and *E. coli*. We anticipate this new regulatory element could be used to build RNA based networks that implement biological control strategies, or to tune the threshold response of a biosensor.

3.2 Main Text

A central goal of synthetic biology is to create libraries of modular parts for the precise and predictable control of gene expression. Libraries of synthetic parts exhibiting a wide range of functionality have now been built and characterized (McLaughlin et al. 2018; Madsen et al. 2016), and great progress has been made in combining these parts into

---

This work is being prepared for publication in *ACS Synthetic Biology* as a Technical Note.

sophisticated genetic networks (Nielsen et al. 2016). While many networks have been constructed with existing parts, it is still highly desirable to develop strategies to tune the function of regulatory parts to accomplish or optimize a desired network function.

Of the many types of genetic parts developed, RNA parts have emerged as a particularly versatile toolset for controlling gene expression and engineering genetic networks. In particular, RNA parts have recently been developed that regulate many aspects of gene expression including transcription (Lucks et al. 2011; Chappell et al. 2015; Qi et al. 2013), translation (Green et al. 2014; Pardee et al. 2016), and mRNA degradation (Carrier & Keasling 1999; Carothers et al. 2011; Pfleger et al. 2006). Additionally, RNA-based genetic circuits have many potential advantages including the possibility of leveraging advances in RNA folding algorithms and design rules for part design (Takahashi et al. 2016; Zadeh et al. 2011) and their naturally fast circuit dynamics (Takahashi, Chappell, et al. 2015). While RNA regulators can be tuned through rational design of RNA structures (Chappell et al. 2017; Salis et al. 2009), here we sought to develop a *trans*-acting strategy to use molecular sequestration interactions to interfere with RNA regulators, a strategy which has been successfully used to tune transfer functions for protein regulators (Chen & Arkin 2012) and is beginning to be explored in the context of RNA regulators.

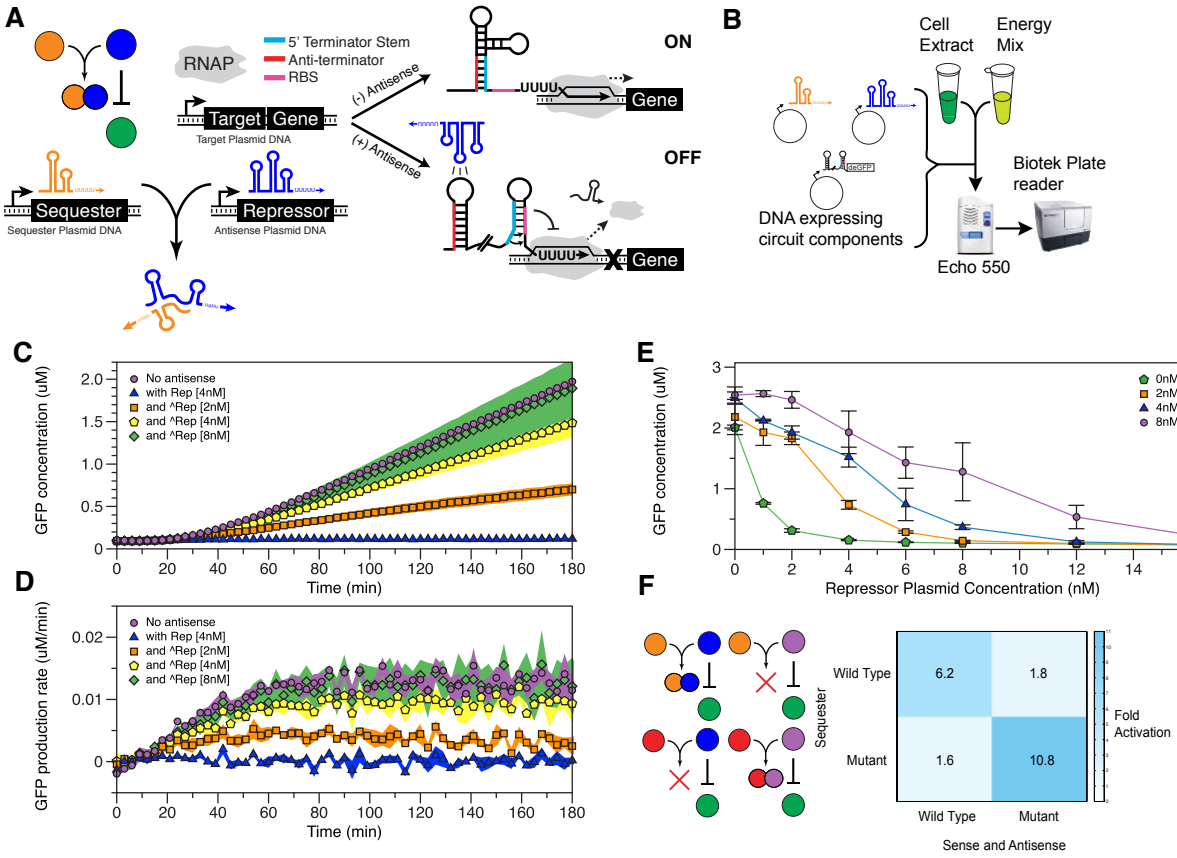


Figure 3.1. RNA sequestration can be used to tune RNA repressor function in cell free transcription-translation (TXTL) reactions. (A) Schematic of the repressor/sequester system. The pT181 target sequence resides in the 5' untranslated region and regulates the expression of a downstream gene. In the absence of repressor RNA (dark blue), the target folds such that the anti-terminator sequence (purple) sequesters the 5' region of the terminator stem (light blue), preventing terminator formation and allowing transcription elongation by RNA polymerase (grey). This structure also contains an exposed ribosome binding site (RBS, red), which allows ribosomes to bind and translate the mRNA. Thus in the absence of repressor RNA the target is transcriptionally and translationally ON. When the repressor RNA is present, it interacts with the anti-terminator, allowing terminator formation, which prevents downstream transcription, and also occludes the RBS inside the

3’ side of the terminator hairpin to prevent ribosome binding. Thus in the presence of repressor RNA the target is transcriptionally and translationally OFF. An RNA sequester (orange) binds to the second hairpin of the RNA repressor and blocks its interaction with the target RNA, leading to a net activation of gene expression. (B) Schematic of the TXTL experiment performed with an Echo 550 for rapid characterization of circuit combinations. TXTL components (cell extract and energy mix) are thawed on ice and then combined with purified DNA expressing circuit elements by acoustic droplet ejection with an Echo 550. The prepared reactions are then moved to the BioTek Plate reader to measure fluorescence over time. (C) Functional time course characterization of repressor sequestration *in* TXTL. GFP fluorescence is measured every 3 minutes for TXTL reactions containing 1nM reporter plasmid, with either no repressor or sequester plasmid (purple), with 4nM of repressor plasmid (blue), and with concentrations of sequester plasmid at 2, 4, and 8nM (orange, yellow, green respectively). Total DNA concentration was kept fixed with a control plasmid. (D) Production rates were calculated from the data in (C). (E) Sequestration shifts the repressor transfer function in TXTL. GFP fluorescence at 3 hours is plotted versus the repressor plasmid concentration. Sequester plasmid concentrations range from 0 to 8nM and cause the repressor transfer function to shift from the left at 0nM (green) to the right at 8nM (purple). (F) Characterization of sequester orthogonality in TXTL. The wildtype and mutant target/repressor pairs are challenged with their cognate sequester and noncognate sequester. Fold activation, calculated by dividing the de-repressed expression by the repressed expression, is only high with cognate pairs indicating low crosstalk.

To develop an RNA sequestration mechanism, we sought to build upon an RNA

repressor that has been shown to be useful for engineering a growing number of RNA circuits (Lucks et al. 2011; Takahashi, Chappell, et al. 2015; Hu et al. 2015) and has been optimized for low circuit leak (Westbrook & Lucks 2017). The repressor is designed from the pT181 attenuation mechanism from the *Staphylococcus aureus* plasmid pT181 (Novick et al. 1989), and is made up of two interacting RNAs: a target RNA and a repressor RNA. The target is an RNA sequence in the 5' untranslated region (5' UTR) of a gene of interest. Alone, the target RNA co-transcriptionally folds into a structure that allows for transcription and translation of the downstream gene. When the repressor RNA is present, its binding to the RNA target causes the 5' UTR to fold into a structure that exposes a transcriptional terminator upstream of the gene of interest during transcription. This terminator represses transcription of the mRNA (Figure 3.1A) by causing the polymerase to ratchet off the DNA, and blocks translation by occluding the ribosome binding site (RBS) within the terminator hairpin. A number of RNA engineering strategies have utilized the pT181 attenuator as a starting point to create RNA genetic networks and gene expression logics (Lucks et al. 2011), build a library of orthogonal transcriptional repressors (Takahashi & Lucks 2013), and most recently invert this mechanism to build RNA transcriptional/translational activators (Chappell et al. 2015; Westbrook & Lucks 2017). Furthermore, a variety of genetic circuits have been constructed with these orthogonal pT181-based regulators including logic gates (Lucks et al. 2011; Chappell et al. 2015), transcriptional cascades (Lucks et al. 2011), single input modules (Takahashi, Chappell, et al. 2015), and negative autoregulation feedback networks (Hu et al. 2018). In this work we tune pT181 repression by introducing an RNA sequester that binds to the second hairpin of the pT181 antisense, blocking its ability to repress its target (Figure 3.1A).

While versatile, it has been difficult to fine tune the regulatory response of the pT181 mechanism using RNA engineering strategies. Therefore, in this work we tune pT181 repression by developing a trans-acting sequestering RNA that is fully complementary to the second hairpin of the pT181 antisense so that it can bind and block its ability to repress its target (Figure 3.1A). We began by evaluating the performance of the designed sequester by testing it's functionality in a cell free transcription-translation (TXTL) expression system using an Echo 550 Liquid Handler which uses focused sound pulses to quickly move small volumes of liquid. The combination of TXTL, which allows rapid characterization of genetic parts and circuits, with the Echo system, makes it possible to characterize many combinations of circuit components in parallel by simply varying the DNA concentrations of individual circuit elements (Santora 2018). Since TXTL characterization data also shows good agreement with *in vivo* data (Takahashi, Hayes, et al. 2015), this setup allowed us to perform allowing rapid prototyping in TXTL to inform follow-up in vivo experiments.

To rapidly characterize repressor and sequester function in TXTL, we used the Echo to titrate antisense and sequester plasmid concentrations using a fixed concentration of reporter plasmid. Four plasmids were used that constitutively expressed the repressor, sequester, Target-GFP, or a control RNA (Figure 3.1B). In each TXTL reaction we added 1nM of the target-GFP reporter plasmid, a varying concentration of the antisense plasmid (0, 1, 2, 4, 6, 8, 12, or 16nM) and a varying concentration of the sequester plasmid (0, 2, 4, 8nM). A control plasmid that expresses a terminator RNA was used to maintain a constant DNA concentration in each reaction. For each combination of plasmids, we then measured the GFP fluorescence over time using a plate reader (Figure 3.1B). Since there is no protein degradation in standard TXTL reactions (Garamella et al. 2016), we use measurements of

both bulk fluorescence over time (Figure 3.1C) as well as fluorescence production rate (Figure 3.1D) to analyze the system dynamics and observe a steady state as a plateau in production rate. Production rates can then be used to calculate % repression (100%-OFF gene expression rate/ON gene expression rate) in TXTL (Takahashi, Chappell, et al. 2015). Without any sequester, we observed 99% repression of GFP production rate which agrees with previous data showing 98% repression in vivo (Westbrook & Lucks 2017). When sequester is added, it is expected bind to the repressor RNA to blocking repression of GFP and lead to a net activation of gene expression. Indeed, with the range of sequester plasmids tested we found that the addition of sequester could tune repression from 74% repression with 2nM of sequester plasmid to 0% repression with 8nM of sequester plasmid (Figure 3.1D). This result shows that sequestration is able to tune repression in TXTL.

We were also able to use this data set to determine how the presence of the sequester shifts the transfer function of the repressor system in TXTL. A transfer function describes the relationship between the input and the output of a system, here input repressor concentration and output gene expression. Transfer functions are often used to characterize genetic parts as they can help predict the response of a system to a given input (Carbonell-Ballestero et al. 2014). The transfer function for the RNA repressor can be seen in Figure 3.1E (green), which shows the expected decrease of GFP output when the repressor plasmid concentration is increased. When a sequester is added, its presence is expected to increase the amount of repressor needed to observe the same level of gene expression, causing an overall shift to the right of the transfer function (Chen & Arkin 2012). As predicted, when increasing concentrations of sequester plasmid are added, the transfer function gradually shifts to the right indicating a higher threshold for switching with a higher concentration of

sequester. This shows that the sequester does shift the transfer function of the repressor interaction in TXTL.

In order for these parts to be used in larger circuit contexts they need to be orthogonal. Libraries of orthogonal RNA parts have been developed for both activators (Chappell et al. 2017) and repressors (Takahashi & Lucks 2013). To demonstrate that sequesters can be designed to maintain orthogonality, we specifically challenged two versions of repressor/target pairs with their corresponding cognate and non-cognate repressors sequesters (Figure 3.1F). Indeed, we observed minimal crosstalk between the non-cognate pairs indicating that these sequesters are orthogonal and could be used in circuitry together.

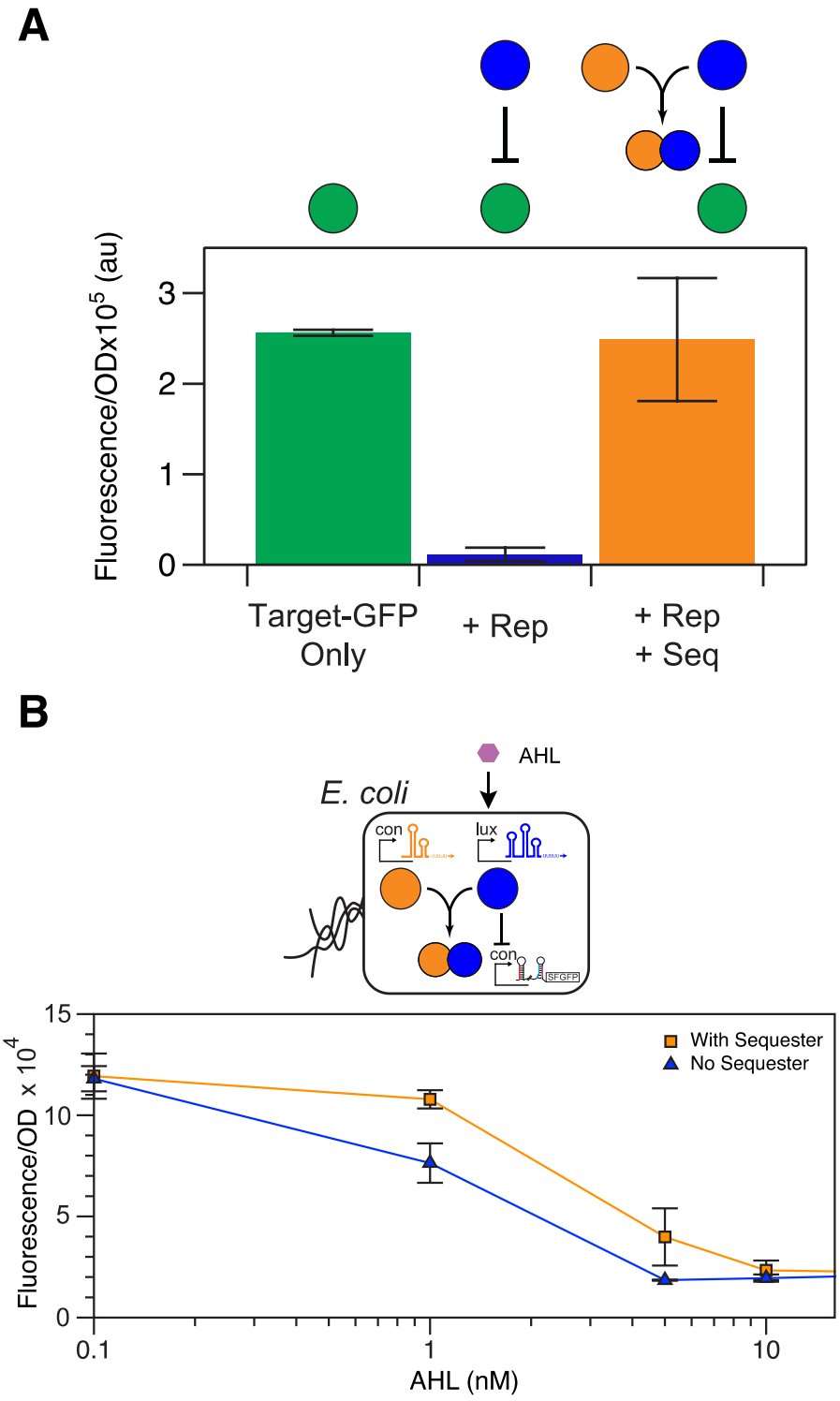


Figure 3.2. RNA sequestration can be used to tune RNA repressor function *in vivo*. (A) Functional characterization of the repressor/sequester system *in vivo*. A BioTek plate reader

was used to measure bulk fluorescence of *E. coli* TG1 cells transformed with plasmids constitutively expressing the target-SFGFP on a low copy plasmid (green), the repressor on a medium copy plasmid (blue), and the sequester on a high copy plasmid (orange). Error bars represent standard deviations of at least three biological replicates. Cartoons indicate which parts are present for each of the three conditions. (B) The repressor transfer function *in vivo* with and without sequester. The repressor is under control of the Lux promoter on a high copy plasmid, which is activated by the addition of AHL to the media. Bulk fluorescence was measured for different concentrations of AHL while cells that were constitutively expressing the reporter plasmid, and either a plasmid expressing a control RNA (blue) or a sequester RNA (orange) on a medium copy plasmid. The error bars indicate the standard deviation from three biological replicates.

After confirming functionality in TXTL we sought to test our sequestration strategy in *E. coli*.. To characterize function in *E. coli*, each RNA part was placed on a separate plasmid. The sequester was cloned onto a high copy plasmid, the antisense onto a medium copy plasmid, and the target onto a low copy plasmid. The plasmids were transformed into *E. coli* TG1 cells, and individual colonies were picked, grown overnight, sub-cultured into minimal media and grown until logarithmic growth was reached. Fluorescence was measured for each culture using a plate reader. When only the target-GFP (Figure 3.2A, green) is present we see high GFP expression. As expected, the repressor gives 96% repression (Figure 3.2A, blue) and when the sequester is added, GFP expression is restored to 3% repression (Figure 3.2A orange). Notably, we observed an increase in variability when the sequester is present, which could be due to the increased burden of RNA expressed in the system.

We also tested whether the sequester could shift the repressor transfer function in *E. coli.* To test this, we cloned the antisense on a high copy plasmid following a pLux promoter, which enabled us to use acyl homoserine lactone (AHL) added to the media to induce expression. The sequester was moved to a medium copy plasmid under constitutive expression and the target remained on a low copy plasmid. We transformed *E. coli* TG1 cells with the circuit plasmids (with and without the sequester) and the cells were grown in varying concentrations of AHL added to the media to induce antisense expression. For both transfer functions in the presence and absence of sequester, we observed AHL induced SFGFP repression (Figure 3.2B). When the sequester is expressed, it shifts the threshold of AHL needed to repress SFGFP to higher levels, confirming that transfer function shifting can be implemented by our designed sequester in *E. coli*.

This work demonstrates an effective strategy to tune repression and shift transfer functions in TXTL and *E. coli.* Previous attempts at RNA sequestration have been successfully used to tune CRISPR repression (Lee et al. 2016) and STAR activation (Lee et al. 2018). Our system provides an RNA only method of tuning repression with high efficiency and low circuit leak. We anticipate that this tool could be used to tune an RNA-mediated biosensor by shifting it's sensing threshold, or as a method to calculate error with respect to a reference RNA concentration for biological control strategies (Agrawal et al. 2018).

3.3 Materials and Methods

TXTL Extract and Buffer Preparation.

Cell extract and reaction buffer were prepared according to Garamella et al. (Garamella et al. 2016)

TXTL experiments.

TX-TL buffer and extract tubes were thawed on ice for approximately 20 min. Separate reaction tubes were prepared with combinations of DNA representing a given circuit condition. Appropriate volumes of DNA, buffer, and extract were calculated using a custom spreadsheet developed by Sun et al. (Sun et al. 2013) and modified to fit the experiment. Buffer and extract were mixed together and then added to each tube of DNA according to the previously published protocol. Each TX-TL reaction mixture (10 μL each) was transferred to a 384-well plate (Nunc 142761), covered with a plate seal (Nunc 232701), and placed on a Biotek SynergyH1m plate reader. We note that special care is needed when pipetting to avoid air bubbles, which can interfere with fluorescence measurements. Temperature was controlled at 29°C. GFP fluorescence was measured (485 nm excitation, 520 emission) every 5 min.

Strains, growth medium, and In Vivo gene expression.

All experiments were performed in *E. coli* strain TG1. Plasmid combinations were transformed into chemically competent *E. coli* TG1 cells, plated on Difco LB+Agar plates containing antibiotics and incubated overnight at 37 °C. Plates were taken out of the incubator and left at room temperature for approximately 9 h. Three colonies were picked and used to inoculate 300 μL of LB containing antibiotics at the concentrations above in a 2 mL 96-well block (Costar 3960), and grown approximately 17 h overnight at 37 °C at 1,000 rpm in a Labnet Vortemp 56 benchtop shaker. 50 μL from each of the fresh cultures was removed from the 96-well block and transferred to a 96-well plate (Costar 3631) containing

50 µL of phosphate buffered saline (PBS). SFGFP fluorescence (FL, 485 nm excitation, 520 nm emission) and optical density (OD, 600 nm) were then measured using a Biotek Synergy H1m plate reader.

3.4 Funding

This work was supported by the Defence Advanced Research Projects Agency (Contract HR0011-16-C01-34).

3.5 References

Agrawal, D.K. et al., 2018. Mathematical Modeling of RNA-Based Architectures for Closed Loop Control of Gene Expression. ACS synthetic biology, p.acssynbio.8b00040.

Carbonell-Ballestero, M. et al., 2014. A bottom-up characterization of transfer functions for synthetic biology designs: lessons from enzymology. Nucleic acids research, 42(22), pp.14060–14069.

Carothers, J.M. et al., 2011. Model-driven engineering of RNA devices to quantitatively program gene expression. Science (New York, N.Y.), 334(6063), pp.1716–1719.

Carrier, T.A. & Keasling, J.D., 1999. Library of synthetic 5' secondary structures to manipulate mRNA stability in Escherichia coli. Biotechnology progress, 15(1), pp.58–64.

Chappell, J. et al., 2017. Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. Nature Communications, 8(1), p.795.

Chappell, J., Takahashi, M.K. & Lucks, J.B., 2015. Creating small transcription activating RNAs. Nature chemical biology, 11(3), pp.214–220.

Chen, D. & Arkin, A.P., 2012. Sequestration-based bistability enables tuning of the switching boundaries and design of a latch. Molecular Systems Biology, 8(1), p.620.

Garamella, J. et al., 2016. The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. ACS synthetic biology, 5(4), pp.344–355.

Green, A.A. et al., 2014. Toehold switches: de-novo-designed regulators of gene expression. Cell, 159(4), pp.925–939.

Hu, C.Y. et al., 2018. Engineering a Functional small RNA Negative Autoregulation Network with Model-guided Design. ACS synthetic biology, p.acssynbio.7b00440.

Hu, C.Y., Varner, J.D. & Lucks, J.B., 2015. Generating Effective Models and Parameters for RNA Genetic Circuits. ACS synthetic biology, 4(8), pp.914–926.

Lee, Y.J. et al., 2016. Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. Nucleic acids research, 44(5), pp.2462–2473.

Lee, Y.J., Kim, S.-J. & Moon, T.S., 2018. Multilevel Regulation of Bacterial Gene Expression with the Combined STAR and Antisense RNA System. ACS synthetic biology, 7(3), pp.853–865.

Lucks, J.B. et al., 2011. Versatile RNA-sensing transcriptional regulators for engineering genetic networks. Proceedings of the National Academy of Sciences of the United States of America, 108(21), pp.8617–8622.

Madsen, C. et al., 2016. The SBOL Stack: A Platform for Storing, Publishing, and Sharing Synthetic Biology Designs. ACS synthetic biology, 5(6), pp.487–497.

McLaughlin, J.A. et al., 2018. SynBioHub: A Standards-Enabled Design Repository for Synthetic Biology. ACS synthetic biology, 7(2), pp.682–688.

Nielsen, A.A.K. et al., 2016. Genetic circuit design automation. Science (New York, N.Y.), 352(6281), pp.aac7341–aac7341.

Novick, R.P. et al., 1989. pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. Cell, 59(2), pp.395–404.

Pardee, K. et al., 2016. Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. Cell, 165(5), pp.1255–1266.

Pfleger, B.F. et al., 2006. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. Nature biotechnology, 24(8), pp.1027–1032.

Qi, L.S. et al., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152(5), pp.1173–1183.

Salis, H.M., Mirsky, E.A. & Voigt, C.A., 2009. Automated design of synthetic ribosome binding sites to control protein expression. Nature biotechnology, 27(10), pp.946–950.

Santora, A., 2018. LABCYTE®: Application Note | High-throughput Microliter-Sized Cell-Free Transcription-Translation Reactions for Synthetic Biology Applications Using the Echo® 550 Liquid Handler | Marshall et al., pp.1–6.

Sun, Z.Z. et al., 2013. Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. Journal of visualized experiments : JoVE, (79), p.e50762.

Takahashi, M.K. & Lucks, J.B., 2013. A modular strategy for engineering orthogonal chimeric RNA transcription regulators. Nucleic acids research, 41(15), pp.7577–7588.

Takahashi, M.K. et al., 2016. Using in-cell SHAPE-Seq and simulations to probe structure-function design principles of RNA transcriptional regulators. RNA (New York, N.Y.), 22(6), pp.920–933.

Takahashi, M.K., Chappell, J., et al., 2015. Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. ACS synthetic biology, 4(5), pp.503–515.

Takahashi, M.K., Hayes, C.A., et al., 2015. Characterizing and prototyping genetic networks with cell-free transcription–translation reactions. Methods, 86, pp.60–72.

Westbrook, A.M. & Lucks, J.B., 2017. Achieving large dynamic range control of gene expression with a compact RNA transcription-translation regulator. Nucleic acids research, 45(9), pp.5614–5624.

Zadeh, J.N. et al., 2011. NUPACK: Analysis and design of nucleic acid systems. Journal of computational chemistry, 32(1), pp.170–173.

CHAPTER 4

DISTINCT TIMESCALES OF RNA REGULATORS ENABLE THE CONSTRUCTION OF A GENETIC PULSE GENERATOR

4.1 Abstract

To build complex genetic networks with predictable behaviors, synthetic biologists use libraries of modular parts that can be characterized in isolation and assembled together to create programmable higher-order functions. Characterization experiments and computational models for gene regulatory parts operating in isolation are routinely employed to predict the dynamics of interconnected parts and guide the construction of new synthetic devices. Here, we individually characterize two modes of RNA-based transcriptional regulation, using small transcription activating RNAs (STARs) and CRISPR interference (CRISPRi), and show how their distinct regulatory timescales can be used to engineer a composed feedforward loop that creates a pulse of gene expression. We use a cell-free transcription-translation system (TXTL) to rapidly characterize the system, and we apply Bayesian inference to extract kinetic parameters for an ODE-based mechanistic model. We then demonstrate in simulation and verify with TXTL experiments that the simultaneous regulation of a single gene target with STARs and CRISPRi leads to a pulse of gene expression. Our results suggest the modularity of the two regulators in an integrated genetic circuit, and we anticipate that construction and modeling frameworks that can leverage this modularity will become increasingly important as synthetic circuits increase in complexity.

This work has been submitted for publication in Biotechnology and Bioengineering (Westbrook et al. 2018).

4.2 Introduction

An important goal of synthetic biology is the development of rational methods for precise temporal control of gene expression, which is necessary to achieve sophisticated dynamic functions in engineered cells (Chappell, Westbrook, Verosloff & Lucks 2017b; Lucks et al. 2011; Green et al. 2014; Carrier & Keasling 1999). Towards this broad goal, libraries of synthetic regulatory parts have been developed to give synthetic biologists control over distinct levels of gene expression (Takahashi, Chappell, et al. 2015). In order to create more complex networks, these parts need to be modular and composable (Chappell et al. 2015), performing their function within the network with minimal undesired interactions. RNA provides a powerful platform to achieve this.

RNA-based regulators have become increasingly popular for building libraries of synthetic parts to orthogonally control many aspects of gene expression (Qi et al. 2013). RNA transcriptional regulators are particularly interesting because they can regulate RNA synthesis as a function of an RNA input and thus can be used to create genetic circuitry that propagates signals on the RNA level (Chappell, Westbrook, Verosloff & Lucks 2017b). These circuits have many potential advantages over protein-based circuits, including the ability to leverage RNA-folding algorithms and high-throughput structure determination to optimize regulatory part folding and function (Chappell, Westbrook, Verosloff & Lucks 2017a), reduced metabolic load for the host (Qi et al. 2013), and rapid signal propagation due to their fast degradation rates (Alon 2013).

Here we focus on building a simple genetic network by combining two modes of RNA-based transcriptional regulation: using small transcription activating RNAs (STARs) (Goentoro et al. 2009) and clustered regularly interspaced short palindromic repeats

(CRISPR) interference (CRISPRi) (Entus et al. 2007; Cheng et al. 2017; Barone et al. 2017). STARs activate gene expression through an interaction with a sequence specific target RNA. The target RNA resides in the 5' UTR of the gene of interest and folds into a transcriptional terminator that halts transcription by causing the polymerase to fall off of the DNA complex before the downstream gene. When present, the activating RNA – called the STAR – binds to the target RNA to prevent terminator formation, thus allowing downstream transcription to turn gene expression ON (Figure 4.1A). Libraries of orthogonal STARs have been built and shown to work in many contexts, including within genomic DNA to reprogram cellular phenotypes, and to control multiple genes within a metabolic pathway (VG 2015; Liao et al. 2017; Nielsen et al. 2016).

CRISPRi is a method of transcriptional repression that relies on targeting a catalytically dead Cas9 (dCas9) nuclease to a gene (Alon 2013; Del Vecchio & Murray 2017). Targeting is dictated by a guide RNA (gRNA) with a 20-bp segment complementarity to the sequence of interest. Here we use the *Streptococcus pyrogenes* Cas9 that targets sequences flanked by a 3' NGG PAM. Binding of the dCas9:gRNA ribonucleoprotein complex to DNA can either block polymerase binding if the targeted region is near a promoter or halt transcription elongation if the targeted region is within a gene. Orthogonal gRNAs can be designed to independently regulate multiple genes or to integrate signals for genetic circuits such as logic gates (Hu et al. 2015; O'Brien et al. 2012). In nature, gRNAs are produced by RNase III cleavage of dsRNA formed by the binding of a trans-activating crRNA (tracrRNA) to complementary sequences in a transcribed CRISPR RNA (Hu et al. 2015). The resulting processed crRNA binds to Cas9 (or dCas9) to form an active ribonucleoprotein complex (Figure 4.1B). CRISPRi works efficiently using either gRNAs produced by the

processing of crRNA/tracrRNA duplexes or using single-guide RNAs (sgRNAs) which fuse the tracrRNA and crRNA to mimic the processed form using a single molecule (Subsoontorn et al. 2012; Brown & Sethna 2003). In this work we use separate crRNA and tracrRNA because they represent the natural form of the gRNA as it is expressed in bacteria, and they also add to an additional time delay in the CRISPRi regulation due to the kinetics of pairing between the RNAs.

One difference between STAR and CRISPRi mechanisms is the timescale on which the regulation occurs. STARs rely on one co-transcriptional RNA-RNA interaction that results in transcription activation typically within minutes (Subsoontorn et al. 2012), while CRISPRi requires the formation of an RNA-protein repressor complex before DNA binding to DNA for repression, which has been shown to take on the order of one hour for regulation to occur (Chappell, Westbrook, Verosloff & Lucks 2017b). This timescale difference between these two opposing modes of gene regulation thus creates an intriguing possibility to use STARs and CRISPRi to engineer a network that produces a pulse of gene expression, similar to the incoherent type-1 feedforward loop (I1-FFL) (Garamella et al. 2016).

The I1-FFL is a common network motif in natural bacterial networks (Garamella et al. 2016) and has received much interest due to its ability to produce a pulse of gene expression (Sun et al. 2013) and accelerate the response time (Mangan et al. 2006). I1-FFLs have also been used to implement band-pass filters (Entus et al. 2007; Kaplan et al. 2008), fold-change detection (Goentoro et al. 2009), biosensing (Barone et al. 2017), and noise buffering (Osella et al. 2011). An I1-FFL consists of an activator X that activates a gene Z and simultaneously its repressor, Y (Figure 4.1C). It can produce a pulse of gene Z

expression because the activation reaction is triggered immediately by X, while the dominating repression occurs with a delay due to the presence of the intermediate component Y (Mangan & Alon 2003). Here, we exploit STARs to induce rapid activation of gene expression, and CRISPRi to achieve delayed repression due to the slow assembly of the gRNA-dCas9 complex. We expect that, when combined, these two RNA-based regulatory mechanisms will operate on timescales that are sufficiently different to yield a transient pulse of gene expression (Figure 4.1D). While our design is not an I1-FFL by a strict definition, it accomplishes the same general behaviour and should produce a pulse of gene expression by exploiting the regulatory timescale differences to cause the delayed repression of Z after fast activation.

A challenge in interconnecting molecular components characterized in isolation is that unexpected interactions between species and resource competition can affect the predicted operation of the composed system, as demonstrated previously (Qian et al. 2017). Reaction rates can be affected by possible crosstalk between the components and the relative abundance of RNA species and dCas9, which are subject to biological noise and circuit complexity (Mishra et al. 2014), thus making the prediction of the integrated construct dynamics necessary and challenging. To address these challenges, we use an interdisciplinary approach that combines cell-free experiments and mathematical modeling.

Mathematical models have gained popularity in guiding the construction and characterization of dynamic molecular systems, given their cost-effectiveness and efficiency as compared to experiments (VG 2015; Liao et al. 2017; Nielsen et al. 2016; Hu et al. 2015). Ordinary differential equations (ODEs) are an effective tool to model molecular

reaction networks, gene expression in protein-based genetic network systems (Alon 2013; Del Vecchio & Murray 2017), and small RNA transcriptional circuits (Hu et al. 2015; O'Brien et al. 2012). ODEs are particularly suitable to model and parameterize cell-free reactions, where initial concentration of chemical species can be accurately controlled. In order to rapidly characterize the STAR and CRISPRi reactions we developed ODE models based on experiments performed with TXTL, an *E. coli* cell-free transcription-translation platform (Sun et al. 2013). TXTL experiments have been successfully combined with mathematical models to parameterize and understand RNA circuits (Hu et al. 2015; Hu et al. 2018; Agrawal et al. 2018). TXTL is ideal for prototyping genetic circuit dynamics because it is quick and easy to use, requires minimal cloning, and shows good agreement with *in vivo* data (Takahashi, Hayes, et al. 2015), and recently it was used to to characterize CRISPR nucleases and guide RNAs (Marshall et al. 2018). Additionally, TXTL also allows for experiments that would otherwise be difficult to perform *in vivo* by giving direct control over component concentrations and enabling circuit optimization and flexibility when designing experiments to fit model parameters.

Here, we start by using TXTL to verify that the STAR and CRISPRi present sufficiently distinct regulatory timescales. Then, we build ODE models for the STAR and CRISPRi pathways in isolation, and we perform systematic TXTL experiments to parameterize and validate the models. We find that when the models are composed to build the IFFL circuit, they predict the expected pulse generation. We conclude with experiments showing that, when connected together to regulate the same promoter, the candidate STAR-CRISPRi pulse generator circuit yields a pulse in target gene expression, and that the composed models can quantitatively capture the pulse generator behavior.

Our results demonstrate that the combination of modeling and experiments in a simplified TXTL environment is an effective approach to prototyping biological dynamic circuits for control of gene expression. Most importantly, our results indicate that RNA regulators characterized in isolation can be combined in more complex circuits without loss of performance when interconnected, making them modular and composable components for dynamic synthetic circuits.

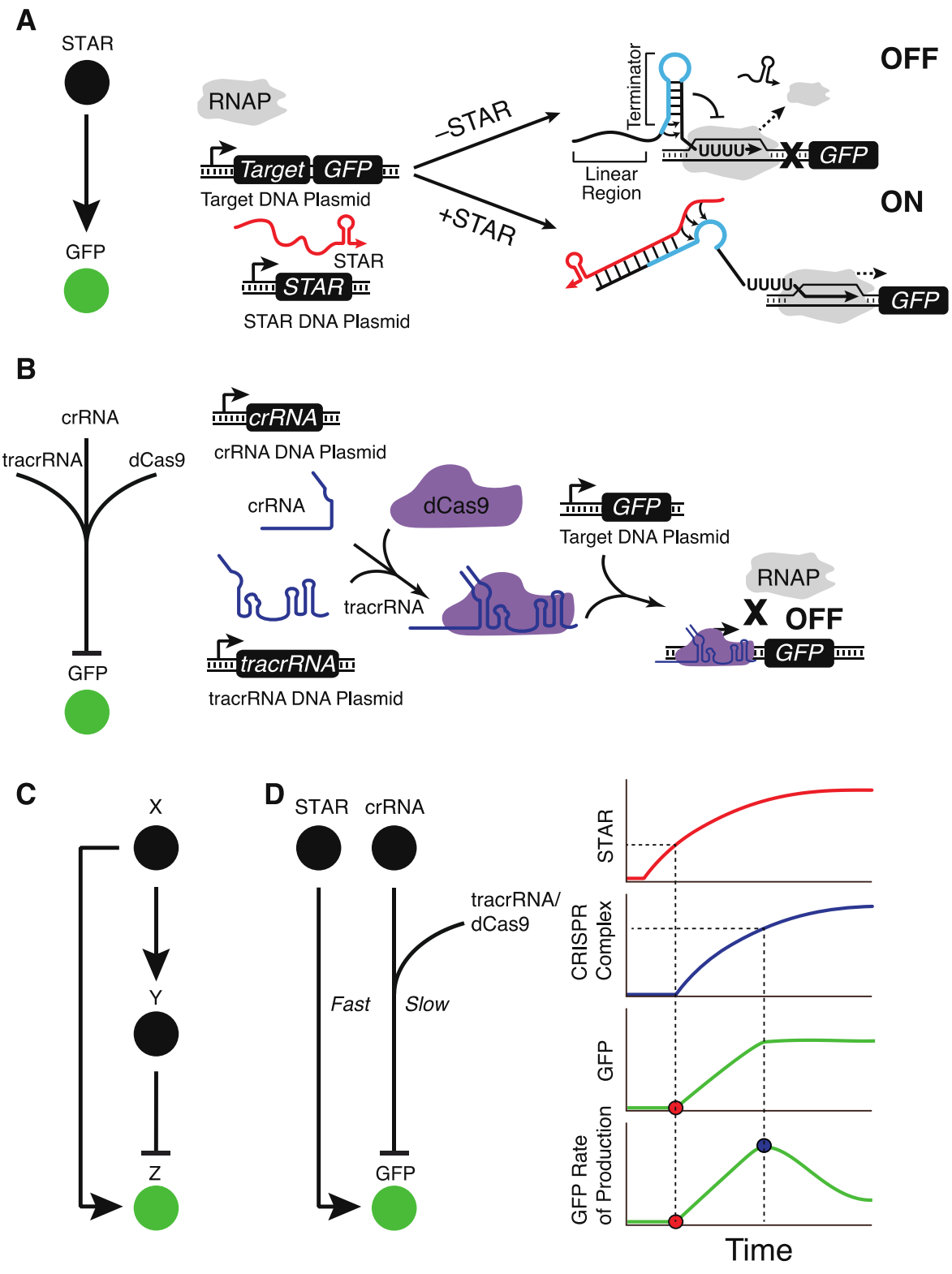

A
STAR
GFP
RNAP
Target
GFP
Target DNA Plasmid
STAR
STAR
STAR DNA Plasmid
-STAR
+STAR
Terminator
Linear Region
UUUU
X
GFP
OFF
ON
B
crRNA
tracrRNA
dCas9
GFP
crRNA
crRNA DNA Plasmid
dCas9
tracrRNA
tracrRNA DNA Plasmid
GFP
Target DNA Plasmid
RNAP
X
OFF
GFP
C
X
Y
Z
D
STAR
crRNA
tracrRNA/
dCas9
Fast
Slow
GFP
STAR
CRISPR Complex
GFP
GFP Rate of Production
Time

Figure 4.1. Architecture of a type 1 incoherent feedforward loop (I1-FFL) composed of STAR activation and CRISPRi repression. (A) Small transcription activating RNA (STAR) mechanism. The target RNA sequence folds into a transcriptional terminator (blue) that causes RNA polymerase to ratchet off the DNA complex and halt transcription upstream of the gene (gene OFF). When present, a STAR (red) binds to both the linear region and the 5' half of the terminator hairpin (blue) of the target RNA, preventing terminator formation and allowing transcription elongation of the gene (gene ON). (B) CRISPR interference mechanism. The crRNA, tracrRNA, and dCas9 bind to form the CRISPR complex that specifically binds to a DNA sequence encoded by the crRNA sequence. When bound the CRISPR complex either blocks transcription initiation or transcription elongation. (C) The I1-FFL motif consists of three parts. An activator X activates expression of Z and its repressor, Y. (D) The pulse generator circuit works by taking advantage of fast STAR activation and slow CRISPRi repression. STAR activates GFP expression immediately while the crRNA/tracrRNA/dCas9 formation causes a delay before finally repressing GFP expression. In TXTL there is no protein degradation, so this causes a pulse in the rate of GFP production.

4.3 Results

4.3.1 Pre-incubation experiments confirm the expected STAR/CRISPRi timescale difference

We first sought to verify the timescale difference between STAR and CRISPRi regulation expected from previous studies (Chappell, Westbrook, Verosloff & Lucks 2017b; Chappell et al. 2015). To do this, we designed experiments that isolated the kinetic processes of each

mechanism. We transcribed RNA components and allowed folding and complex formation with previously synthesized dCas9 before assessing regulatory function, to isolate only the timescale of the regulatory mechanism. When performing a typical TXTL experiment, all DNA is added to the reaction at t = 0 and gene expression is measured over the course of a few hours. Inherent to this experimental design is a delay due to the transcription of RNA regulator parts, which must first be transcribed before they can perform their function. In order to isolate the timescale of the regulatory event, we incubated a plasmid expressing each RNA regulatory part alone for 2 hours, essentially allowing the TXTL reaction to synthesize RNA regulators before being assessed for function. We then mixed pre-incubated reactions with reporter DNA and characterized the response time of the system. In this way, we removed the timescale needed for regulatory RNA synthesis and instead focused the characterization experiment on the relevant timescales of action for each regulator.

The STAR system only has one trans-acting RNA, so we incubated a plasmid expressing the STAR RNA or a plasmid expressing a non-functional control RNA in TXTL for 2 hours. We then added DNA encoding the p70a-Target-GFP plasmid to this reaction mixture at t = 0 and began measuring fluorescence over time. We observed detectable STAR activation of gene expression ~20 minutes after the addition of the GFP plasmid (Figure 4.2A) and STAR activation as determined by the GFP production rate reached 70% of the steady state after 35 minutes, where the steady state was computed from an exponential fit (Figure 4.2B) as described in Supplementary Note B1.

We anticipated that regulation of gene expression by the dCas9 complex would take significantly longer than the STAR activation, given previous observations suggesting that

gRNA loading onto dCas9 takes on the order of ~1hr in the presence of non-specific RNAs (Mekler et al. 2016). As the CRISPRi system requires a crRNA, tracrRNA, and dCas9, a more sophisticated experiment was required to characterize the regulatory timescale. Specifically, we sought to determine the timescale for crRNA-tracrRNA-dCas9 complex assembly required for the dCas9 complex to repress gene expression. To quantitatively estimate this timescale, we incubated the DNA encoding each RNA component in all combinations of alone, together, and in TXTL already containing dCas9 for 2 hours (Supplementary Figure B1) and then combined them into a final reaction with DNA encoding the p70a-GFP plasmid before began measurement. For clarity, we only show two conditions in Figure 4.2C: all alone or all together in TXTL containing dCas9. When incubated separately, we expect all components to be present at high concentrations at the beginning of the measurement but no CRISPRi repression complex will have formed yet. The complex will begin forming when the measurement starts. When incubated together, we expect the CRISPRi complex to have already formed and be present at high concentrations. Comparing these two conditions indicates the time it takes for the crRNA-tracrRNA-dCas9 complex to form and then repress (Figure 4.2C). However, when incubated separately, the complex was slower to repress gene expression, and did not achieve full gene repression until 55 minutes after addition of the DNA reporter construct (Figure 4.2D). This large difference in response times reveals that the crRNA-tracrRNA-dCas9 complex takes on the order of 55 minutes to fully form and perform its function in TXTL, which is similar to previous research (Mekler et al. 2016).

Taken together, these results indicate that there is a timescale difference between STAR activation (70% of the steady state production rate seen after 35 minutes) and

CRISPRi repression (30% peak production rate seen after 55 minutes) due to the extra steps required for the crRNA-tracrRNA-dCas9 complex assembly as opposed to the direct RNA-RNA interactions of the STAR mechanism. These timescale differences could therefore be exploited to construct a simple network architecture that produces a pulse of gene expression.

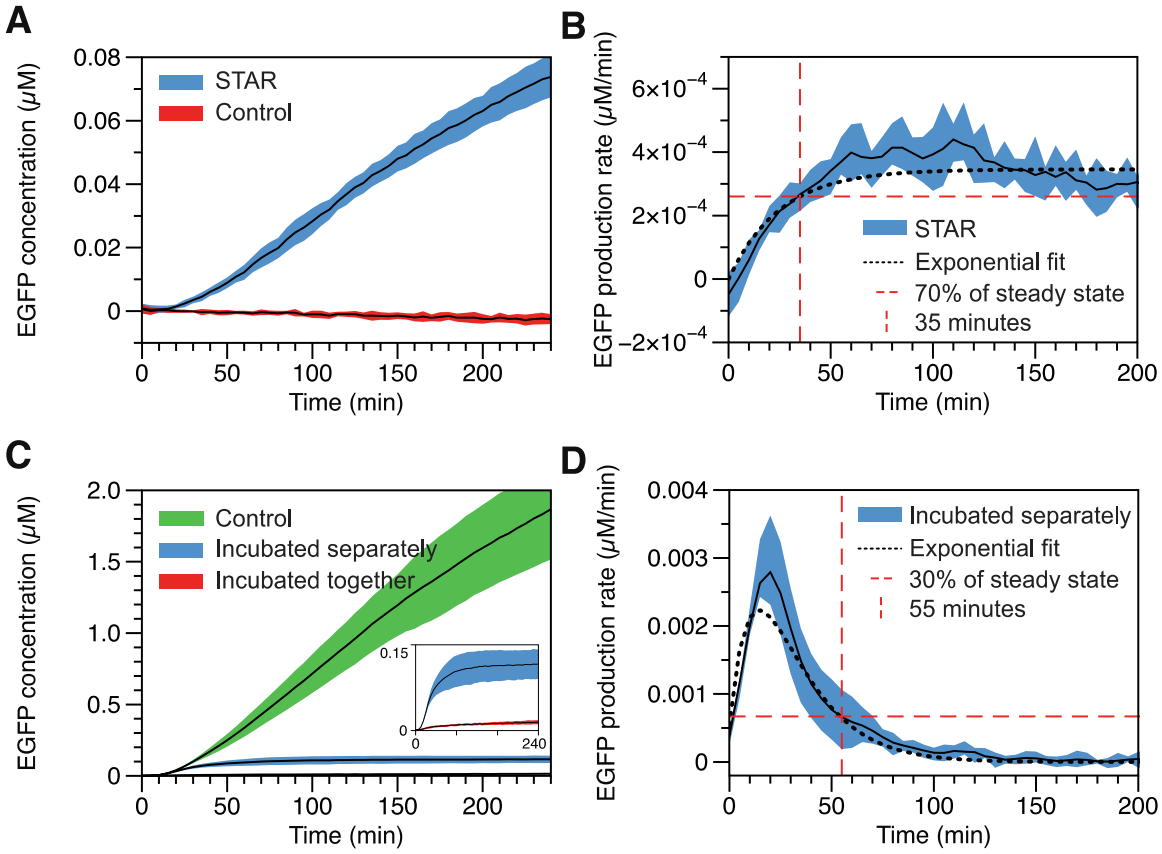


Figure 4.2. Pre-incubation experiments indicate that STAR activation is faster than dCas9-based repression. (A) Functional time course characterization of GFP expression when STAR is pre-incubated (blue) or a non-function control is pre-incubated (red). The timescale of STAR activation is on the order of 20 minutes after reporter DNA is added to the reaction. (B) The production rate of GFP expression for the STAR pre-incubation experiment (blue). The GFP production rate reached 70% of max as determined by the

exponential fit (dotted black line) at 35 minutes. (C) Functional time course characterization of the CRISPRi response when parts are incubated together (red) or separately (blue) in comparison to unrepressed expression (green). The timescale of pre-incubated CRISPR repression is much faster than when the parts are incubated separately, suggesting that the dCas9 loading time adds a significant delay to the system. The inset shows the two repressed states. Data for all pre-incubation combinations of CRISPRi parts is shown in Supplementary Figure B6. (D) The production rate of GFP expression for the CRISPRi pre-incubation experiment (blue). The GFP production rate reached 30% of its peak as determined by the exponential fit to the derivative (dotted black line) at 55 minutes.

4.3.2 STAR and CRISPRi model derivation

After verifying the timescale difference with our pre-incubation experiments, we then sought to construct mathematical models for the STAR and the CRISPRi systems respectively, to computationally test our hypothesis and guide the design of the circuit, before conducting further experiments. We used ordinary differential equations to model the rate-of-change of each molecular concentration, as a result of coupled kinetic reactions (Figure 4.3).

We modelled the STAR activation as a one-step reaction, where STAR binds to the free promoter P_y directly to achieve transcription activation, at rate β_s, mimicking the fact that STAR activation only requires RNA-RNA interactions. This is an approximation that coarse-grains the details of how the small RNA modifies target RNA structure to activate transcription, but it is justified based on similar simplifying assumptions made in previous

work modeling RNA transcriptional repressors (Hu et al. 2015; Hu et al. 2018). In parallel, we modelled CRISPR-Cas9 complex formation as a two-step reaction process. As part of this process, the tracrRNA and the crRNA bind to form the gRNA at rate γ_1. The gRNA can then bind to dCas9 to form the active repressor complex at rate γ_2. Since dCas9 dissociation rates are extremely low with no mismatches (Boyle et al. 2017), we assumed the formation of the CRISPR-Cas9 complex and its binding (at rate ω) to the free promoter P_y to form the repressed promoter P_y^- to be irreversible. While capturing the key reactions in the CRISPR-Cas9 formation, the model coarse-grains the detailed dynamics of how crRNA, tracrRNA and dCas9 interact with each other and interferes transcription. To enable a direct comparison between the STAR and CRISPRi regulation pathway, we used a first-order kinetic reaction to model the STAR activation, instead of the Hill-type function used in Hu et al (Hu et al. 2015).

In the STAR system, reporter p70a-Target-GFP mRNA (M) is only produced when p70a-Target-GFP (P_y) is activated (i.e. bound to STAR at rate β_s), at rate α_m, while in the CRISPRi system, M is only produced from the free promoter p70a-GFP (for simplicity and for later use in the combined model, this is also denoted by P_y), at rate α_m. The GFP translational initiation, elongation, and maturation were modelled following prior work (Hu et al. 2015), and the mature GFP (G_m) is compared to the experimental measurement.

In addition to the transcriptional rates above, each RNA species has a degradation rate and each protein species has a translation rate. Specifically, α_s, α_{cr}, α_{tr}, δ_s, δ_{cr}, δ_{tr}, and δ_g are the transcriptional and degradation rates of STAR, crRNA, tracrRNA, and gRNA respectively; δ_m is the degradation rate of GFP mRNA, M; K_i is the translation initiation rate, K_e is the translation elongation rate, and α_{gm} is the GFP maturation rate. P_y^{tot} is the total

amount of reporter promoters, M_i is the translationally initialized mRNA, and G is the immature GFP protein. We note that no protein degradation rate is included because proteins do not degrade in TXTL unless degradation tags are included (Garamella et al. 2016). and there is no translation rate for dCas9 because extracts were made from *E. coli* cells expressing dCas9 (Marshall et al. 2018).

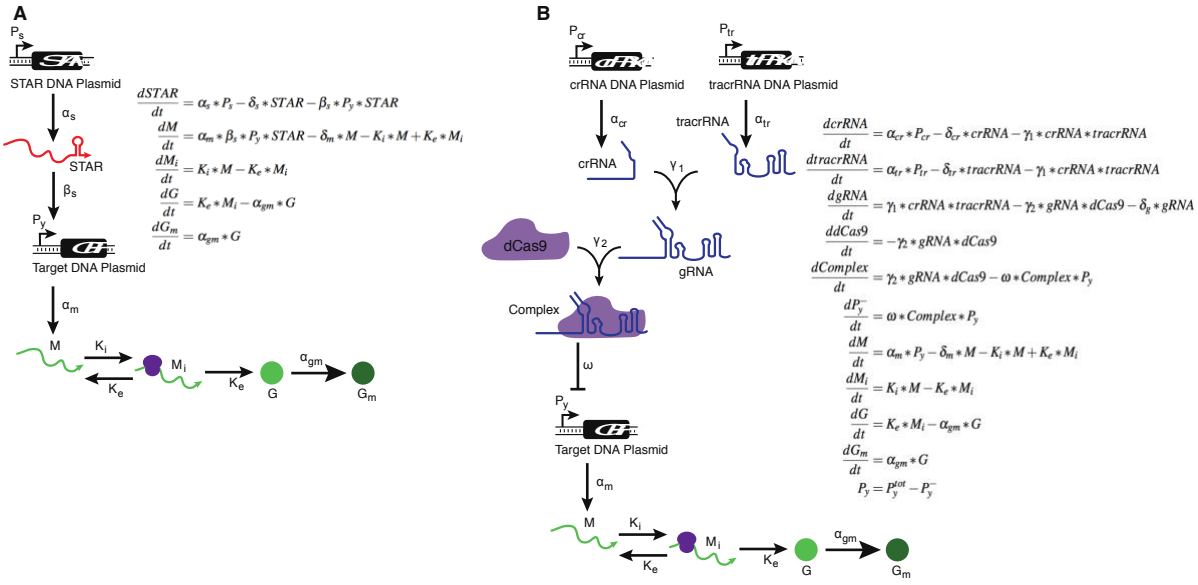


Figure 4.3. Separate STAR and CRISPRi models with the corresponding topology. The STAR activation is modelled as a one-step binding at rate β_s, to the free output promoter P_y, to enable expression of GFP mRNA M. The CRISPRi repression is modelled as a two-step reaction, where formation of active repressor complex happens before it binds to the free output promoter P_y to form the repressed P_y^-, and GFP is only expressed from the free P_y promoter. For simplicity, the degradation rates of the RNA species are modelled but not shown in the topology. In both models, mature GFP protein G_m is compared to experimental measurements. All the STAR, M, Mi, G, G_m, crRNA, tracrRNA, gRNA, Complex, and P_y^- are initiated with concentration 0 nM. The initial free P_y plasmid was 0.5 nM, and dCas9 concentration was estimated to be 35 nM based on previous experimental measurement.

4.3.3 Model parameterization

With the separate STAR and CRISPRi models, our next step was to extract suitable kinetic parameters to construct a combined model for reliable predictions. To achieve this, we adopted a Bayesian inference parameterization approach (Subsoontorn et al. 2012; Brown & Sethna 2003) to fit parameters for STARs and CRISPRi separately. Specifically, we used three sets of the STAR activation experiments (Full experimental data shown in Supplementary Figure B2) to train our model for the three STAR-related kinetic parameters: α_s, δ_s, and β_s. We also used three sets of the CRISPRi repression experiments (Full experimental data is shown in Supplementary Figure B3) to train our model for the eight CRISPRi-related kinetic parameters: α_{cr}, α_{tr}, δ_{cr}, δ_{tr}, δ_g, γ_1, γ_2, and ω. As crRNA and tracrRNA were transcribed from the same promoter in our experiments, we assumed that they share the same transcription rate, and we set $\alpha_{cr} = \alpha_{tr}$ in the fitting. The five reporter GFP-related parameters (α_m, δ_m, K_i, K_e, and α_{mg}) were also fitted for both STAR and CRISPRi.

For both STAR and CRISPRi experiments, we initiated our fitting with 10 different initial guesses that were evenly spaced in the admissible parameter intervals that were inferred from previous publications (Supplementary Table B1) (Hu et al. 2015). To fit the 8 parameters in the STAR model, we conducted 105,000 iterations of parameter updates to seek convergence, and to fit the 12 parameters in the CRISPRi model, we conducted 210,000 iterations. The probability of accepting parameter set *i* from parameter set *j* was set according to the following (Subsoontorn et al. 2012):

$$P(i \leftarrow j) = \begin{cases} e^{-\frac{E_i - E_j}{T}}, & \text{if } \Delta E = E_i - E_j \geq 0 \\ 1, & \text{otherwise} \end{cases} \qquad (1)$$

with T = 0.125 (= $2\sigma^2$), and σ is the estimated measurement error. The cost function *E* is defined as the cumulated point-wise squared prediction-measurement error for each experiment cycle:

$$E = \sum_{t=t_0}^{t_f} \left[Prediction(t) - Measurement(t) \right]^2 \qquad (2)$$

The parameter set that gave the lowest cost function *E* across all the fittings was deemed as the best fitted parameter set. The corresponding simulations are plotted in Figure 4.4A and B against the experimental measurement. The comparisons between predictions and data demonstrate that models trained with the Bayesian inference approach were able to reproduce the dynamics of the STAR and the CRISPRi system under various conditions. To understand the distribution of each parameter, we ranked all the sampled parameter sets (i.e. 10 x 105,000 and 210,000 sets of parameters for the STAR and CRISPRi fitting respectively) with respect to the corresponding value of the cost function *E*. Figure 4.4C shows the parameter distribution of the first 1000 sets of parameters that gave the lowest fitting error *E*. Note that the five GFP-related parameters shown in Figure 4.4C were fitted from the STAR activation experiments, for demonstration. The values of the best fitting parameters are given in Supplementary Table B1.

Interestingly, while some CRISPRi-related parameters have a relatively wide distribution, we see limited variation in the repressor formation-related parameters such as ω for the plotted 1,000 fitted parameter values. This observation suggests that the repressor formation kinetics dominate the accuracy of the CRISPRi regulation process. On the other hand, all three STAR-related parameters displayed a relatively wide distribution, which suggests the existence of multiple optimal solutions for the fitting. This might be due

to our simplification of the STAR activation mechanism and/or limited experimental conditions (e.g. initial concentrations), such that a wide range of parameter values can fit well the model. Note that fewer reaction steps and experimental conditions lead to fewer constraints for the parameterization. The corresponding fitted parameter distribution and correlation are given in Supplementary Figures B4 and B5.

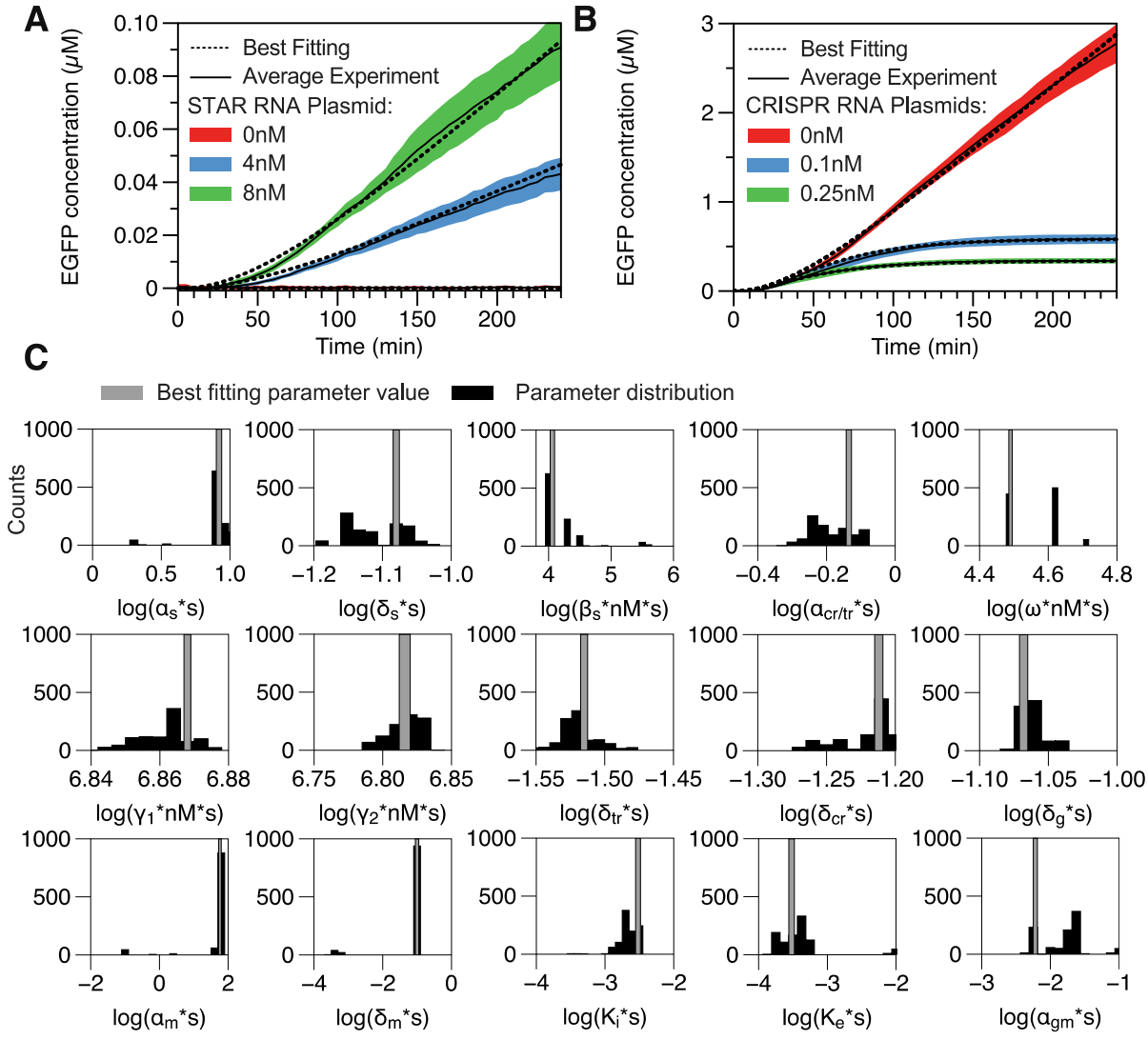


Figure 4.4. Model parameterization with separate STAR and CRISPRi experiments. (A) Comparison of best fitted simulation to the STAR experiments for three conditions: high activation with 8 nM of STAR plasmid (green plots, 8 nM STAR), moderate activation with 4

nM of STAR plasmid (blue plots, 4 nM STAR), and no activation with no STAR plasmid (red plots, STAR OFF). (B) Comparison of the best-fitted simulation to the CRISPRi parameterization experiments for three conditions: no repression with no crRNA or tracrRNA (red plots, 0nM CRISPR RNA), moderate repression with 0.1 nM crRNA and tracrRNA plasmid (blue plots, 0.1 nM CRISPRi RNA), and complete repression with 0.25 nM crRNA and tracrRNA plasmid (green plots, 0.25 CRISPRi RNA). (C) Histogram of parameters obtained from 1000 samples that gave the lowest fitting error within the pool of 10 x 105000 and 10 x 210000 fitting rounds for the STAR and CRISPRi system respectively. Grey bar indicates the location of the parameter value that gave the best fitting. Note, all the kinetic parameters are scaled to be dimensionless before taking their log values in the histogram plots.

4.3.4 Pulse generator modeling and experimental verification

After parameterizing the separate STAR and CRISPRi models, we then combined them to build the pulse generator model by introducing a competition for P_y promoter binding between STAR and CRISPRi (Figure 4.5). In the combined pulse generator model, a free promoter P_y can either bind to CRISPRi to form a repressed state or to STAR to form an activated state for gene expression. Once P_y is bound to the CRISPRi complex, it becomes unavailable for STAR activation. To simulate the model, all output promoter copies were initiated in the free state (unbound), with a fixed concentration to mimic conditions used in the CRISPRi characterization experiments.

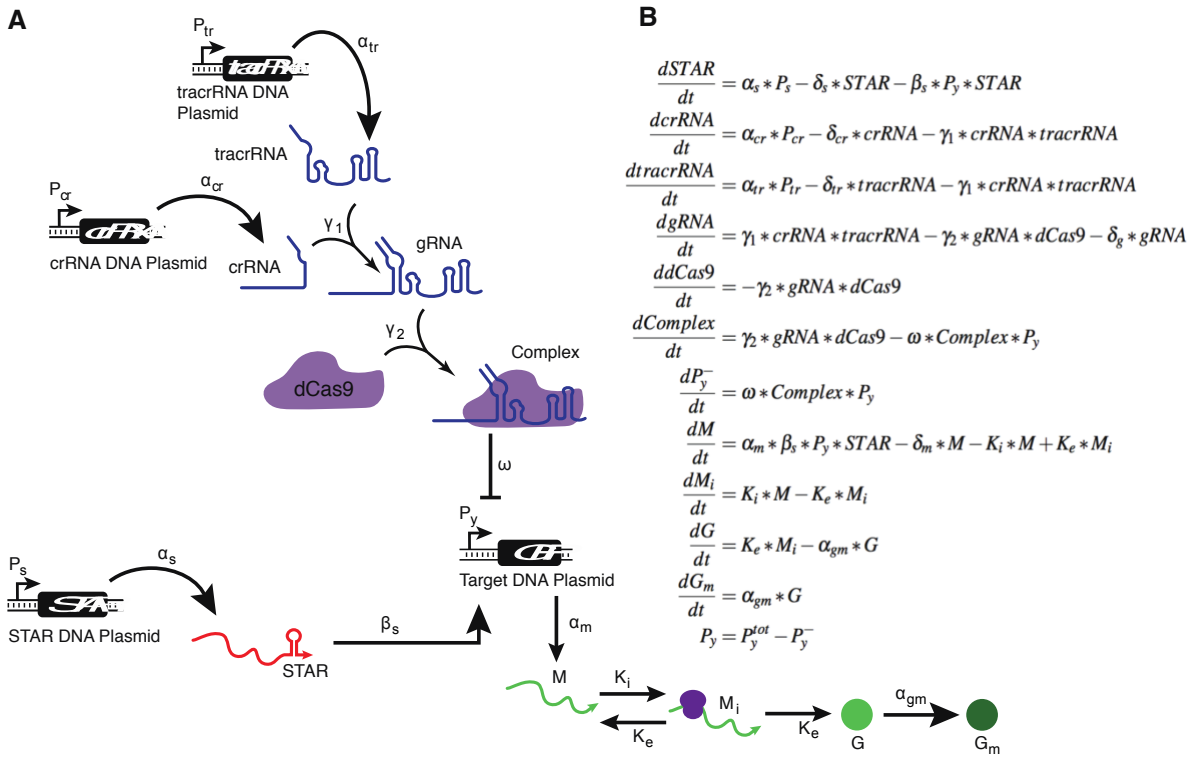


Figure 4.5. Topology of the pulse generator model. The separate STAR and CRISPRi model are combined by introducing a competition for P_y binding through the repressor formation in dP_y^-/dt and the activation in dM/dt equations. Once CRISPRi repressor complex binds to P_y to form repressed state P_y^-, it can no longer be activated for expression. P_y and P_y^- follows mass balance with a total initial concentration of P_y^{tot}. For simplicity, the degradation rates of the RNA species are modelled but not shown in the topology. All the STAR, M, Mi, G, G_m, crRNA, tracrRNA, gRNA, Complex, and P_y^- are initiated with concentration 0 nM. The initial free P_y plasmid was 0.5 nM, and the dCas9 concentration was estimated to be 35 nM based on previous experimental measurement.

We then used the combined model to test if a pulse could be generated in the production rate of the target gene. Instead of using one best fitted parameter set, we

decided to combine the set of best fit from each of the 10 Bayesian fittings for both the STAR and the CRISPRi regulator experiments, obtaining 100 sets of parameters (10 STAR x 10 CRISPRi) to generate predictions of the pulse generator behaviour. This is because a mismatch between our model prediction and the pulse generator behaviour could be caused by the fact that multiple optimal parameters exist for each individual regulator model (Figure 4.4), so the combination of the very best fits might not give the most accurate prediction for the interconnected circuit. The procedure to generate combinations of best fits is summarized in Figure 4.6A. Prediction with the best STAR and CRISPRi separately fitted parameters demonstrated a plateau in the GFP concentration (dashed black plot in Figure 4.6B), and a pulse in the production rate (dashed black plot in Figure 4.6C), and indeed all 100 parameter combinations suggested a pulse in the production rate (Supplementary Figure B6). Given these observations, we expect the integrated pulse generator to function robustly and also to produce a pulse in experiments.

We then performed a TXTL experiment that combined both the STAR and the CRISPRi systems. As in the separate CRISPRi experiment, we added 0.25 nM crRNA, 0.25 nM tracrRNA, and 0.5 nM p70a-Target-GFP. Since the STAR ON expression level is significantly lower than that of the CRISPRi system (Figure 4.4), we doubled the amount of STAR plasmid used in the separate STAR experiment from 8 nM to 16 nM in the combined system, to mitigate this difference. After the addition of all DNAs, we immediately began measuring fluorescent GFP expression (see Supplementary Figure B7 for complete experimental data). As predicted, the experiments also demonstrated a plateau in GFP expression level (Figure 4.6B, red), and a pulse in the production rate (Figure 4.6C, red). We only see a pulse in the production rate because TXTL has negligible protein degradation

(Sun et al. 2013). If performed *in vivo*, we would expect a pulse in concentration rather than production rate. We then quantified the prediction accuracy by defining the prediction error in the same way as the cost function in Eqn. 2 to study the possible changes in the model parameters caused by the combination. The log-based prediction errors are summarized in the heat map in Figure 4.6D.

One interesting observation is that the best prediction (solid black plot in Figure 4.6B and C) was not achieved by the set of the best-fitted parameters (dashed black plot). The best-fitted parameter set predicted a higher steady-state concentration in GFP and a taller pulse in the production rate, as compared to the averaged experimental measurement (solid red plot in Figure 4.6B and C) and the best prediction. Indeed, the best fitted and the best prediction parameters were from the same STAR (same row in Figure 4.6D) but a different CRISPRi fitting trial (different column in Figure 4.6D). The values of the best prediction parameters are given in Supplementary Table B1. This observation suggested that the coupling may affect the CRISPRi dynamics, such that the set of parameters fit best the separate experiments but underpredicted the repressor formation rate, which lead to a higher predicted GFP expression level. A detailed parameter-to-parameter comparison between the best prediction and the best-fitted parameters is given in Supplementary Figure B8, to visualize the relative location of each parameter value.

We next asked how well we can fit the STAR/CRISPRi combined model to the experimental measurements, and how that compares to the best prediction with the separately fitted parameters. Again, to seek convergence we conducted 10 Bayesian fittings from different initial guesses, with 210,000 iterations for each fitting (same as in the CRISPRi fitting). The fitting that yielded the lowest fitting error is plotted in Figure 4.6B and

C in green. Surprisingly, the best prediction with the separately fitted parameters slightly outperformed the best fits on the combined model. This could be due to the fact that in the combined model more parameters have to be simultaneously fitted relative to the individual component models, leading the combined model to require a larger number of samplings (i.e. initial guesses and/or iterations) to reach an equally good fit. Indeed, the fitting error comparison in Supplementary Figure B9 suggests that to fit 12 parameters in the CRISPRi model, even more iterations might be needed. Additionally, the best prediction from the separately fitted parameters is similar to the best fits out of 100 fittings, since it is the best prediction from a 10 x 10 best fitted parameter sets. To improve the fitting on the combined model, one can use more initial guesses and increase the number of iterations. Supplementary Figure B9 summarizes the detailed comparison of the accuracy and the error convergence for the STAR, CRISPRi, and pulse generator model fitting, respectively.

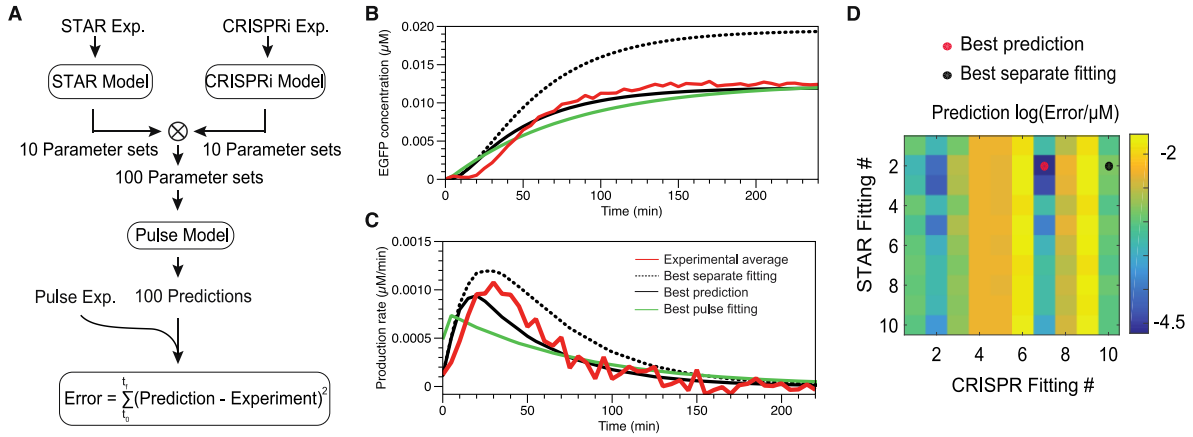


Figure 4.6. Pulse model prediction and experimental verification. (A) Procedure for parameterization and prediction: each of the individual STAR and CRISPRi models were trained with experimental measurements to fit 10 sets of best fitting parameters. These parameters were then combined into 100 sets that were used to predict the dynamics of the pulse model. The predictions were then compared to experimental measurement and

quantified by the squared error between the prediction and the observed trajectories. (B) GFP concentration reached steady state in both the simulation (black and green) and experiments (red) within 240 min, while the best fitted parameter set predicted a higher steady state concentration level (dashed black), and the best prediction from the separately fitted parameters (solid black) gave better accuracy to the best fitting of the pulse model (green); (C) GFP production rate demonstrated a pulse which peaked at around 40 min and dropped when the repression kicked in and RNA degradation took over, in both the simulations (black and green) and the experiments (red). (D) Presentation of the prediction accuracy with the 100 sets of separately fitted parameters indicates the best separately fitted parameter set did not give the best prediction in the combined pulse model. Red dot indicates the location of the best separately fitted parameter sets and the yellow dot indicates the location of the parameter set for the best prediction. Note that they are in the same row (i.e. same STAR fitting trial) but different columns (i.e. different CRISPR fitting trial). Best separate fitting: prediction with parameters that best fit the STAR and CRISPR system individually; best prediction: the best of the 100 predictions with individually fitted parameters; best pulse fitting: best out of the 10 fittings to the pulse experiments.

4.4 Discussion

In this work we have demonstrated an RNA-based pulse generator in TXTL that harnesses the difference in speed between STAR and CRISPRi regulation. This STAR-CRISPRi hybrid construct is able to produce a pulse of gene expression. STAR activation involves a single, fast, co-transcriptional RNA-RNA interaction while CRISPRi requires the slow formation of

an RNA-protein complex leading to a delay before CRISPRi repression sets in. Combined, these mechanisms produce pulse of gene expression caused by the transcription of a few RNA molecules.

There have been a number of synthetic I1-FFLs built using protein regulators (Entus et al. 2007; Cheng et al. 2017; Barone et al. 2017). Recently, we built an RNA-based I1-FFL that uses AHL to activate expression of a STAR RNA that activates expression of mRFP as well as a gRNA and dCas9 that repress mRFP (Chappell, Westbrook, Verosloff & Lucks 2017b). This design relies on an additional RNA cleavage strategy, cascading RNA regulatory events, and slow dCas9 production. Here, we constructed a simpler network that implements the pulse of gene expression of an I1-FFL, but faster and more effectively with a simpler network design.

As synthetic networks grow in complexity, models will be vital for predicting their behaviour and understanding dynamics, as they provide faster assessments of the network as compared to experiments. Here, we constructed and parameterized a coarse-grained mechanistic model and used it to predict the dynamics of the pulse generator network. With the simulation results, we observed possible modularity of the STAR regulator when combined with other structures to form more complicated networks, while the performance of the CRISPRi regulation might be affected, as indicated by the change in the parameter values. However, this observed change in the CRISPRi regulation might be due to several reasons: first, given the limited amount of training data, it could be possible that the CRISPRi parameters were over-fitted on the training data thus giving a non-ideal prediction in the new condition (combined system). Indeed, the complexity, parameterization methods, and experimental noise could all contribute to the accuracy of

the model parameterization. Second, un-modelled (and undesired) coupling of the two regulatory pathways could affect the dynamics; for example, indirect competition for the transcription machinery could reduce transcription rates in a non-homogeneous manner in the two circuits, altering their regulation timescale. Third, the mechanism of the CRISPRi repressor formation might be oversimplified such that intermediate reaction steps were overlooked. For further investigations, we suggest a richer data set under various conditions for model parameterization, and a refined model to encompass more detailed reactions in the system.

Model parameterization can be challenging, especially when obtaining large amount of experimental measurements under various conditions is costly and a stochastic parameterization method is used, which would normally require convergence. The results in this study suggest that, instead of fitting all the parameters simultaneously, fitting part of a combined network separately could also lead to reliable predictions of an integrated structure, especially when the modularity of each component can be maintained. Because fitting parameters of individual modules for use in integrated structures provides a more computationally tractable alternative to comprehensive parameter fitting, we expect this approach to become predominant as synthetic molecular systems become more and more complex.

In summary, we demonstrate a STAR-CRISPRi hybrid pulse generator both with simulation and *in vitro* TXTL experiments; the circuit mimics the architecture and performance of an I1-FFL. We also demonstrated how mathematical modeling can be used to guide and assess the design of biological constructs. We found that parameters fitted from separate models can also accurately predict the performance of the combined

model/construct. We further discussed the importance of sample data, and optimization settings in improving the parametrization. We anticipate the results in this study to provide guideline for future work in the modeling, parameterization, and construction of biological parts made of both STAR and CRISPRi regulators.

4.5 Materials and Methods

Plasmid construction and purification.

Key sequences can be found in Supplementary Table B2. All the plasmids used in this study can be found in Supplementary Table B3. The STAR plasmid and control plasmid were construct pJBL4971 and pJBL002, respectively, from Chappell et al (Chappell, Westbrook, Verosloff & Lucks 2017b). The GFP expression plasmid was p70a-GFP from Garamella et al (Garamella et al. 2016). and the STAR-target plasmid was modified from this plasmid using iPCR. The plasmids expressing crRNA, tracrRNA, and the scrambled crRNA were constructed using Gibson Assembly and iPCR and sequence verified using sanger sequencing. Plasmids were purified using a Qiagen QIAfilter Plasmid Midi Kit (Catalog number: 12243) followed by isopropanol precipitation and eluted with double distilled water.

TXTL Extract and Buffer Preparation.

Cell extract and reaction buffer were prepared according to prior work (Garamella et al. 2016).

TXTL experiments.

TXTL buffer and extract tubes were thawed on ice for approximately 20 min. Separate reaction tubes were prepared with combinations of DNA representing a given circuit condition. Appropriate volumes of DNA, buffer, and extract were calculated using a custom spreadsheet developed by Sun et al (Sun et al. 2013). and modified to fit the experiments. Buffer and extract were mixed together and then added to each tube of DNA according to the previously published protocol. Each TXTL reaction mixture (10 μL each) was transferred to a 384-well plate (Nunc 142761), covered with a plate seal (Nunc 232701), and placed on a Biotek Synergy H1m plate reader. We note that special care is needed when pipetting to avoid air bubbles, which can interfere with fluorescence measurements. Temperature was controlled at 29°C. GFP fluorescence was measured (485 nm excitation, 520 emission) every 5 min. A calibration to EGFP concentration (μM) was performed using a standard curve of pure EGFP (Cell Biolabs STA-201) in order to present measurement data in terms of GFP concentration. Pre-incubation experiments were performed by combining two types of extracts. One extract has dCas9 pre-expressed and the other does not. Each plasmid was incubated in the appropriate extract and buffer for 2 hours before the pre-incubated reactions were combined in equal parts and measurements began.

Modeling.

Equations in Figure 4.3 were solved with MATLAB_R2014b *ode23s* solver to get the simulated GFP concentration for the error calculation (Eqn. 3) in data fitting. Candidate parameters were generated with a uniform distribution within a bounded interval (Supplementary Table B1), using MATLAB random number generation function *rand*. One trial of the Bayesian inference data fitting (i.e. one initial guess, with 105000 iterations)

took about three computational hours on a Macbook Pro with a 2 GHz Intel Core i7 processor. Model in Figure 4.5 was also numerically solved with MATLAB_R2014b *ode23s* function to get predictions for the combined pulse generator.

4.6 Funding

This work was supported by the Defense Advanced Research Projects Agency (contract HR0011-16-C-01-34). The authors declare no conflict of interest.

4.7 Authorship

A. W. and X. T. should be considered co-first authors. A. W., X. T., R. M., C. M., J. C., D. A., M. D., V. N., C. B., J. L., and E. F. all contributed to the conception and design and participated in revision of the manuscript. E. F. and J. L. supervised the work. A. W. and X. T. drafted the manuscript. A. W. performed all data collection. E. F. and X. T. built the model and X. T. carried out the computational work.

4.8 References

Agrawal, D.K. et al., 2018. Mathematical Modeling of RNA-Based Architectures for Closed Loop Control of Gene Expression. ACS synthetic biology, p.acssynbio.8b00040.

Alon, U., 2013. An Introduction to Systems Biology, Chapman & Hall.

Barone, F. et al., 2017. Design and evaluation of an incoherent feed-forward loop for an arsenic biosensor based on standard iGEM parts. Synthetic Biology, 2(1), p.802.

Boyle, E.A. et al., 2017. High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. Proceedings of the National Academy of Sciences of the United States of America, 114(21), pp.5461–5466.

Brown, K.S. & Sethna, J.P., 2003. Statistical mechanical approaches to models with many poorly known parameters. Physical review. E, Statistical, nonlinear, and soft matter physics, 68(2 Pt 1), p.021904.

Carrier, T.A. & Keasling, J.D., 1999. Library of synthetic 5' secondary structures to manipulate mRNA stability in Escherichia coli. Biotechnology progress, 15(1), pp.58–64.

Chappell, J., Takahashi, M.K. & Lucks, J.B., 2015. Creating small transcription activating RNAs. Nature chemical biology, 11(3), pp.214–220.

Chappell, J., Westbrook, A., Verosloff, M. & Lucks, J., 2017a. Computational design of Small Transcription Activating RNAs (STARs) for versatile and dynamic gene regulation. bioRxiv, p.169391.

Chappell, J., Westbrook, A., Verosloff, M. & Lucks, J.B., 2017b. Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. Nature Communications, 8(1), p.795.

Cheng, Y.-Y. et al., 2017. The Timing of Transcriptional Regulation in Synthetic Gene Circuits. ACS synthetic biology, 6(11), pp.1996–2002.

Del Vecchio, D. & Murray, R.M., 2017. Biomolecular Feedback Systems, Princeton University Press.

Entus, R., Aufderheide, B. & Sauro, H.M., 2007. Design and implementation of three incoherent feed-forward motif based biological concentration sensors. Systems and synthetic biology, 1(3), pp.119–128.

Garamella, J. et al., 2016. The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. ACS synthetic biology, 5(4), pp.344–355.

Goentoro, L. et al., 2009. The Incoherent Feedforward Loop Can Provide Fold-Change Detection in Gene Regulation. Molecular Cell, 36(5), pp.894–899.

Green, A.A. et al., 2014. Toehold switches: de-novo-designed regulators of gene expression. Cell, 159(4), pp.925–939.

Hu, C.Y. et al., 2018. Engineering a Functional small RNA Negative Autoregulation Network with Model-guided Design. ACS synthetic biology, p.acssynbio.7b00440.

Hu, C.Y., Varner, J.D. & Lucks, J.B., 2015. Generating Effective Models and Parameters for RNA Genetic Circuits. ACS synthetic biology, 4(8), pp.914–926.

Kaplan, S. et al., 2008. The incoherent feed-forward loop can generate non-monotonic input functions for genes. Molecular Systems Biology, 4(1), p.203.

Liao, C., Blanchard, A.E. & Lu, T., 2017. An integrative circuit-host modelling framework for predicting synthetic gene network behaviours. Nature microbiology, 2(12), pp.1658–1666.

Lucks, J.B. et al., 2011. Versatile RNA-sensing transcriptional regulators for engineering genetic networks. Proceedings of the National Academy of Sciences of the United States of America, 108(21), pp.8617–8622.

Mangan, S. & Alon, U., 2003. Structure and function of the feed-forward loop network motif. Proceedings of the National Academy of Sciences of the United States of America, 100(21), pp.11980–11985.

Mangan, S. et al., 2006. The Incoherent Feed-forward Loop Accelerates the Response-time of the gal System of Escherichia coli. Journal of Molecular Biology, 356(5), pp.1073–1081.

Marshall, R. et al., 2018. Rapid and Scalable Characterization of CRISPR Technologies Using an E. coli Cell-Free Transcription-Translation System. Molecular Cell, 69(1), pp.146–157.e3.

Mekler, V. et al., 2016. Kinetics of the CRISPR-Cas9 effector complex assembly and the role of 3'-terminal segment of guide RNA. Nucleic acids research, 44(6), pp.2837–2845.

Mishra, D. et al., 2014. A load driver device for engineering modularity in biological networks. Nature biotechnology, 32(12), pp.1268–1275.

Nielsen, A.A.K. et al., 2016. Genetic circuit design automation. Science (New York, N.Y.), 352(6281), pp.aac7341–aac7341.

O'Brien, E.L., Itallie, E.V. & Bennett, M.R., 2012. Modeling synthetic gene oscillators. Mathematical biosciences, 236(1), pp.1–15.

Osella, M. et al., 2011. The role of incoherent microRNA-mediated feedforward loops in noise buffering. J. Stelling, ed. PLoS computational biology, 7(3), p.e1001101.

Qi, L.S. et al., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152(5), pp.1173–1183.

Qian, Y. et al., 2017. Resource Competition Shapes the Response of Genetic Circuits. ACS synthetic biology, 6(7), pp.1263–1272.

Subsoontorn, P., Kim, J. & Winfree, E., 2012. Ensemble Bayesian analysis of bistability in a synthetic transcriptional switch. ACS synthetic biology, 1(8), pp.299–316.

Sun, Z.Z. et al., 2013. Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. Journal of visualized experiments : JoVE, (79), p.e50762.

Takahashi, M.K., Chappell, J., et al., 2015. Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. ACS synthetic biology, 4(5), pp.503–515.

Takahashi, M.K., Hayes, C.A., et al., 2015. Characterizing and prototyping genetic networks with cell-free transcription–translation reactions. Methods, 86, pp.60–72.

VG, D., 2015. Synthetic Biology: Computational Modeling Bridging the Gap between In Vitro and In Vivo Reactions. Current Synthetic and Systems Biology, 03(03).

Westbrook, A., Tang, X., Marshall, R., Maxwell, C., Chappell, J., Agrawal, D., Dunlop, M., Noireaux, N., Beisel, C., Lucks, J., Franco, E., 2018. Distinct timescales of RNA regulators enable the construction of a genetic pulse generator. bioRxiv 377572; doi: https://doi.org/10.1101/377572

CHAPTER 5

USING THE TEMPORAL DOMAIN TO ENHANCE CELLULAR COMMUNICATION WITH AN RNA GENETIC CIRCUIT

5.1 Abstract

Current methods for communicating with cells are limited and largely rely on a few inducers to create slow changes in steady state gene expression. We aim to expand on these capabilities by using genetic networks that can decode complex signals. Here we propose two possible circuit designs, a temporal circuit and a circuit that remembers state, that can interpret complex inputs. These circuits receive signal pulses from chemical or light-based inducers and return different gene expression patterns based on the number and type of pulse. Experimental results for the temporal design indicate cell division will need to be artificially slowed for the circuit to function properly. Due to the speed and ease of experimental techniques we suggest light induction is a better choice than chemical induction. Finally, we use a stability analysis to show our circuit design can remember state under some conditions. We anticipate these designs will greatly improve the amount of information that can be passed to cells.

5.2 Introduction

Within the field of synthetic biology, scientists and engineers have reprogrammed organisms with sophisticated synthetic behaviors ranging from production of vital chemicals and drugs to sensing environmental signals. These modified organisms could one day function as single cell sensors and machines to tackle crucial challenges in medicine and sustainability. These desirable cellular behaviors are controlled by patterns of gene expression governed by

networks of genetic regulators. To manipulate these cellular machines, we communicate with them by sending signals that result in changes in gene expression. However, current communication methods are primitive, relying on a few inducers to trigger changes in slow, steady state gene expression. Most genetic circuitry is induced with one of a limited number of steady state chemical signals (IPTG, aTc, arabinose (Voigt 2006), AHL (Basu et al. 2005), or light-based inducers (Levskaya et al. 2005). While progress has been made in engineering synthetic biomolecular parts that can detect signals, most designs focus on logic gates that detect static chemical signals. Temporal information decoding offers a more efficient way to communicate with cells. We aim to improve external cellular communication by engineering cells to detect time-varying signals.

Other synthetic biologists have recently begun to recognize the utility of time-dependent circuits. New innovations have used time varying optical signals to control gene expression (Olson & Tabor 2014). One group has engineered cells that count inducer pulses (Friedland et al. 2009) and another has developed cells that can record order of inducer exposure (Hsiao et al. 2016) opening up the possibility that cellular counting could be used to record environmental histories for applications like microbiome recording and environmental reporting (Kotula et al. 2014). We propose to expand upon this work by developing information processing genetic circuitry that can decode temporal information (Figure 5.1).

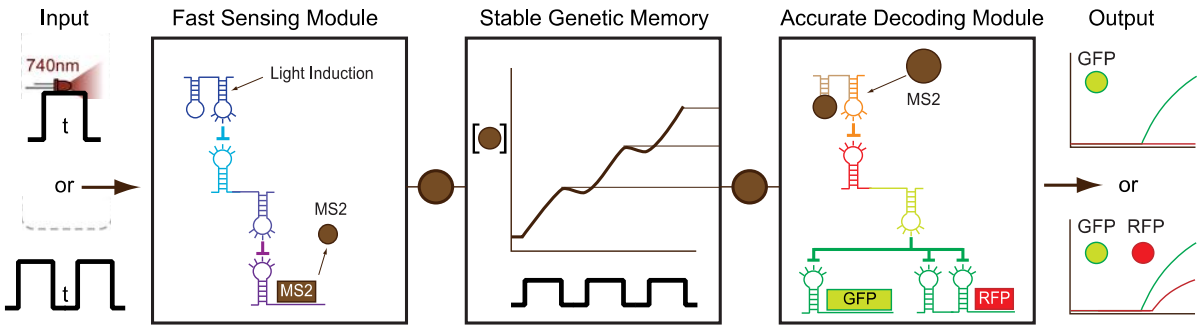


Figure 5.1 One design of an RNA temporal signal decoding circuit. The circuit is composed of two RNA modules buffered by protein-based genetic memory. The sensing module is a light-activated repressor cascade composed of RNA regulators that confers overall activation in the presence of green light. When light is pulsed, the MS2 protein is produced, and acts as a transient genetic memory in the cell. The decoding circuit, which is made from a second RNA repressor cascade, reads MS2 concentration as an input and controls the timing of gene expression as an output. Different configurations of RNA transcription regulators (hairpins) produce different output responses (Levskaya et al. 2005).

Current methods for communicating with cells rely on a limited number of chemical or optical inducers to trigger changes in slow, steady state gene expression. Many natural biological processes are inherently temporal such as developmental transcription networks (Levine & Davidson 2005) and responses to environmental stimuli (Baudrimont & Becskei 2015). Synthetic biologists have begun to recognize the utility of time-dependent circuits. Friedland et al. constructed a synthetic network that could count up to three sequential inducer pulses (Friedland et al. 2009). However, this system does not remember state or distinguish between length and timing of pulses. Pioneering work by Hsiao et al. used unidirectional DNA recombination to record sequences of input events (Hsiao et al. 2016).

This design does remember state but is irreversible. Circuit designs that are fast, robust, reversible, and versatile are needed to address these limitations. We propose to create genetic circuitry that can decode signals in the time domain, alleviating the current bottleneck on external cellular communication.

RNA presents the ideal tool to build this circuitry because its structural and temporal characteristics allow engineers to construct fast, designable genetic networks. Not only does RNA's single stranded structure gives it the flexibility to fold into different forms that determine its function (Chappell et al. 2013), but our increasing understanding of RNA folding and design gives us unique control over its functionality. A few simple circuits have already been built: a transcriptional cascade that confers overall activation and a single input module (SIM) that uses one input to control multiple outputs (Alon 2007; Takahashi et al. 2015). In addition, RNA has been used to construct synthetic NOR, AND, and A-AND NOT-B logic gates, paving the way for genetic computation with RNA (Lucks et al. 2011). In this work we use the pT181 dual control attenuators described in Chapter 2.

Many biological applications require external control of gene expression accomplished through signals sent by molecular inducers. However, there are significant limitations to using chemical signals. Membrane transport processes can limit entrance to the cell and chemicals can be degraded or interfere with cellular metabolism. More specifically to our goals, chemical inducers can be toxic and difficult to pulse in a controlled manner. Optogenetics, or light inducible gene expression offers an attractive alternative to using molecular signals. There has been remarkable progress in using a broad array of light-inducible systems to activate gene expression (Zoltowski et al. 2009; Rockwell et al. 2006).

In synthetic biology, the CcaS/CcaR system (Figure 5.2) (Tabor et al. 2011) has found use in reversibly inducing gene expression and characterizing synthetic parts (Olson et al. 2014).

In is natural context, the cyanobacterial Ccas/CcaR system induces the expression of a phycobilisome-related gene in response to green light. The system consists of a membrane-associated histidine kinase, CcaS, and a response regulator, CcaR. When green light (520nm) is adsorbed by CcaS bound to chromophore phycocyanobilin (PCB), the rate of CcaS autophosphorylation increases. CcaS then transfers its phosphate group to CcaR, which increases transcription of superfolder GFP (SFGFP). Red light (650nm) reverses this process by reducing the rate of CcaS autophosphorylation. The CcaS/CcaR system share a common chromophore (PCB) with the Cph8/OmpR and have been compatibly used to activate gene expression within the same cell (Tabor et al. 2011). The Cph8/OmpR system works similarly to the CcaS/CcaR system except it is activated by far-red light (740nm) and repressed by red light (650nm) with Cph8 being the histidine kinase and OmpR its response regulator. These well characterized optogenetic systems are ideal for inducing genetic circuitry without toxicity in a time-dependent manner.

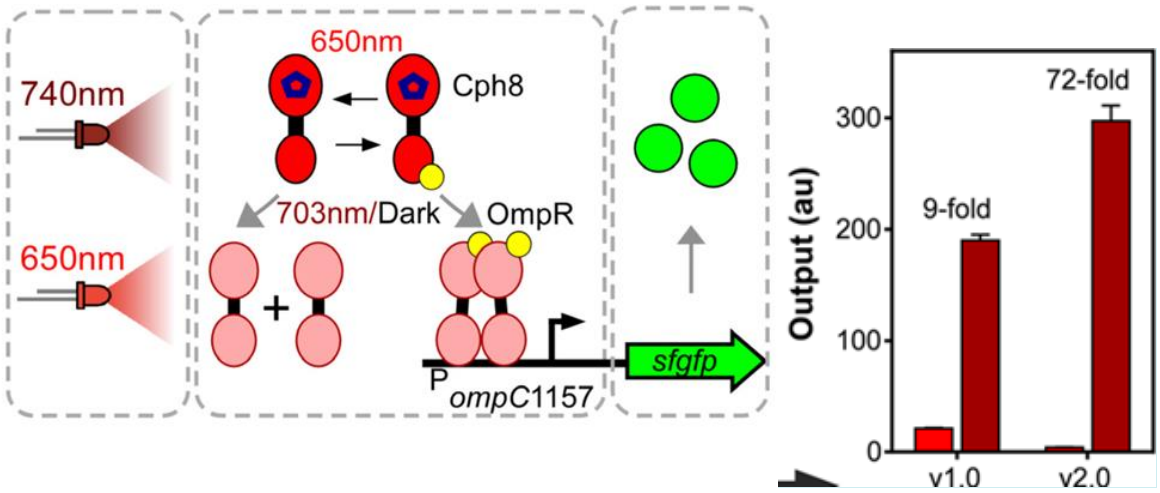


Figure 5.2 The CcaS/CcaR two-component light-switchable system (Schmidl et al. 2014). The system is activated by green (535nm) light and repressed by red (670nm) light. When chromophore phycocyanobilin (PCB, blue pentagons) is bound to CcaS absorption of green light increases the rate of CcaS autophosphorylation, phosphotransfer to CcaR, and transcription of SFGFP. Absportion of red light reverses this process. The system has been engineered to increase fold activation from 6 fold to 117 fold. Figure from Schmidl et al. (Schmidl et al. 2014)

5.3 Results

5.3.1 Designing a temporal signal decoding circuit

Our goal is to build a genetic network that can interpret an input in the time domain and respond with multiple outputs depending on the timing and duration of the input. We have developed two designs that each accomplish this in different ways. The first (Figure 5.1) relies on the hypothesis that RNA degrades much faster than protein, allowing protein to act as genetic memory in an RNA circuit. The second (Figure 5.3) acts as a finite state machine, remembering state with bistability.

The first design is composed of two modules buffered by a transient protein 'memory' (Figure 5.1). The sensing module is a light-activated RNA repressor cascade. When light is pulsed, MS2, a coat protein derived from the MS2 bacteriophage fused to a florescent protein, is produced and acts as genetic memory. The decoding circuit is a single input module (SIM) (Takahashi et al. 2015), made from a repressor cascade, which reads MS2 concentration as an input. The double attenuator in front of red fluorescent protein (RFP) creates a threshold such that at low concentrations of MS2 only green florescent protein (GFP) is produced and

at high concentrations of MS2 both RFP and GFP are produced. After one pulse the cells store an MS2 concentration at which GFP is produced, and after two pulses the cells store an MS2 concentration at which GFP and RFP are produced. Eventually MS2 concentration is reduced through dilution, allowing the circuit to reset over time.

The second design is composed of two connected toggle switches (Figure 5.3). The circuit responds to one or two pulses of green light and is reset by red light. A toggle switch (Figure 5.3 left), similar to an electrical SR latch, is two connected repressors. When the correct conditions are met, a toggle switch exhibits bistability or two states in which one repressor is expressed and the other repressed. One repressor controls GFP such that in one state GFP is high and in the other GFP is low. In the full circuit one toggle switch controls GFP and the other controls RFP. The toggle switches are cascaded such that one pulse causes the first toggle switch to enter a high state and express GFP and a second pulse activates the second toggle switch causing GFP to be repressed and RFP to be expressed. Both toggle switches are turned OFF with the reset, red light.

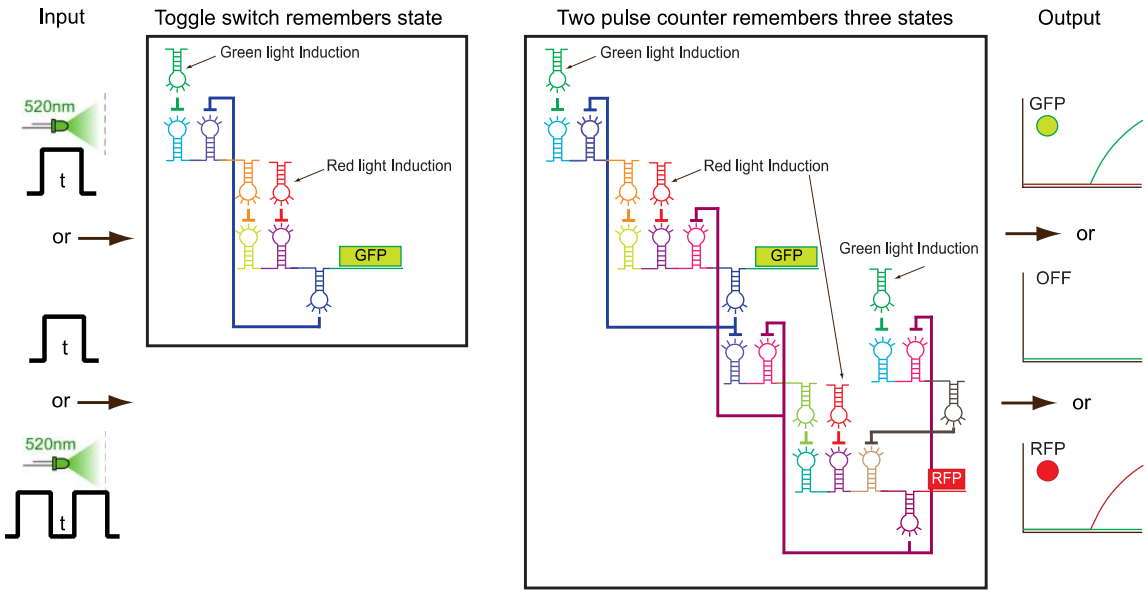


Figure 5.3 Temporal signal decoding circuit design #2. The first design is shown in Figure 5.1. The second design is constructed from two toggle switches and uses green light as a set signal and red light as a reset signal. GFP is produced in response to one green light pulse and RFP in response to two green light pulses. Red light reset the circuit to the OFF state. This circuit remembers state with bistability.

5.3.2 The sensing circuit responds to pulses of theophylline.

We first sought to test the sensing circuit proposed in circuit design #1 (Figure 5.1). We began by testing the module (Figure 5.4A) with theophylline activation instead of light activation. The sensing module is a three level RNA repressor cascade that responds to pulses of theophylline to create pulses of an output protein over time. This output protein is an MS2-RFP fusion, because the MS2 protein is the next input for the planned decoding circuitry. We tested the dynamic response to theophylline pulses in the context of MS2-RFP and verified that we can produce the desired rising saw tooth pattern.

We performed time course experiments on E. coli cultures that contained the dual control sensing module plasmids expressing MS2-RFP. We tested the circuit by incubating overnight in LB media, followed by dilution into M9 supplemented media and incubation for four hours, and then another dilution into fresh M9 and four hour incubation before beginning the measurements to mitigate observed toxicity of theophylline (Figure 5.4B). As expected, when theophylline was pulsed we observed a saw tooth pattern that consisted of a rise in MS2-RFP expression when theophylline was present, and a decrease in MS2-RFP expression when theophylline was absent. Importantly we saw that the second pulse of theophylline could sustain increased MS2-RFP expression.

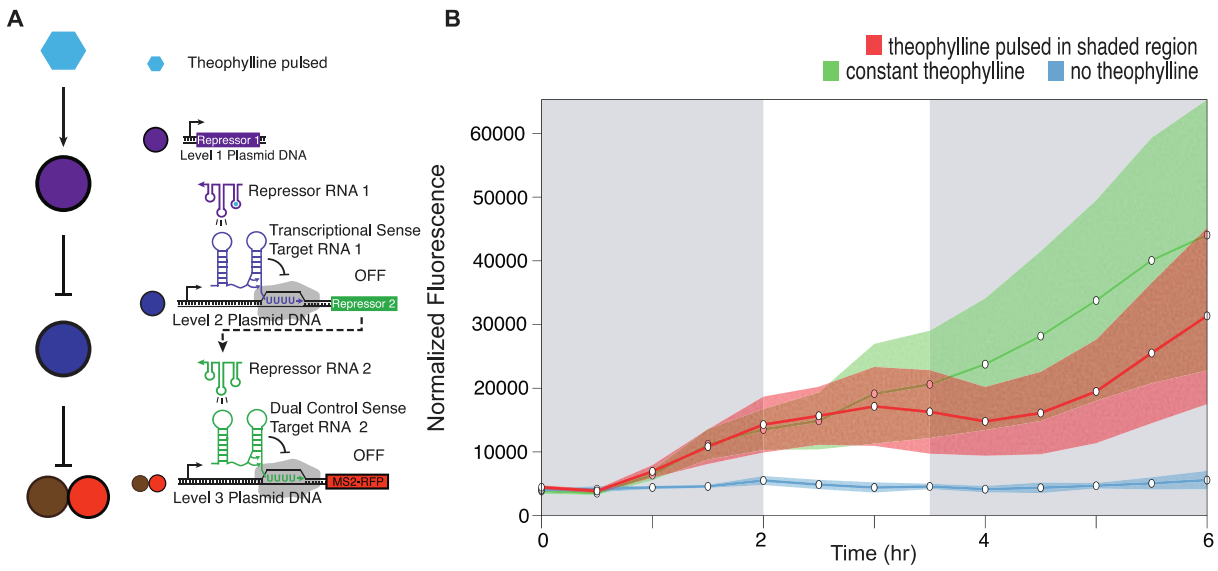


Figure 5.4 The sensing module produces the desired saw-tooth production of GFP and MS2-RFP in response to pulses of theophylline, confirming its desired function. (A) Schematic of the basic sensing circuit. This cascade is identical to Figure 8A except MS2-RFP is produced instead of GFP. (B) Pulse response of the MS2-RFP repressor cascade. Theophylline is added at t=0 to both the constant theophylline (green) and pulsed theophylline (red) cultures. At t=1.5 hours all cultures are spun down at 6,000g for 10 minutes and then resuspended in fresh M9. Theophylline is immediately added to the constant theophylline condition. The t=2.0 hours measurement is taken immediately following resuspension in the correct media. The washing occurs between t=1.5 hours and t=2 hours. At t=3.5 hours theophylline is added again to the pulse condition (red).

5.3.3 The decoding circuit responds to MS2 concentration.

Next we tested the decoding module's response to MS2 concentration. As proposed, the decoding module is a three level RNA repressor cascade that responds to different levels of the MS2 protein to stimulate the expression of different output proteins based on the MS2

concentration sensed. Specifically, this cascade consists of an antisense RNA fused to an MS2 aptamer (Mutalik et al. 2012) in such a way that it is only functional when MS2 is bound. In the presence of MS2, the repression caused by this antisense RNA is inverted through the RNA cascade to produce the desired protein. We constructed and tested the basic functioning of this MS2-sensitive cascade architecture by replacing the theophylline sensitive antisense RNA of the sensing module with the MS2-sensitive antisense RNA and adding pLac controlled MS2-RFP to the top level of the cascade such that MS2-RFP could be induced with IPTG.

To characterize the decoding module, we performed steady state experiments on E. coli cultures that contained the decoding module plasmids (Figure 5.5A). After incubating overnight in LB media, the cultures were diluted into M9 supplemented media, induced with varying concentrations of IPTG, and incubated for four hours. After the four-hour incubation, the cultures were sampled and measured for GFP fluorescence, RFP fluorescence, and culture OD (Figure 5.5B). As expected, increasing IPTG concentration caused MS2-RFP to be produced and this in turn caused increased GFP expression. We also observed a significant leak in GFP expression in the low IPTG conditions. Overall, the decoding module works as expected: increasing MS2 production yields increased GFP expression.

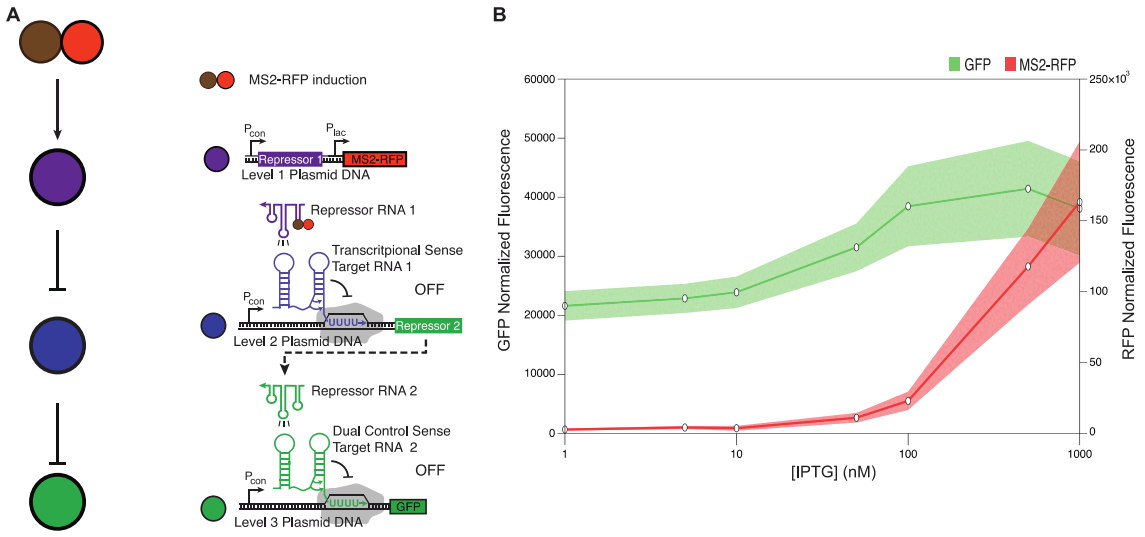


Figure 5.5 The decoding circuit architecture senses the MS2 protein and produces GFP. (A) Schematic of the basic decoding architecture, an MS2 activated repressor cascade. This cascade is very similar to Figure 5.4A except the circuit senses MS2 instead of theophylline. To make MS2 inducible, it was included in the level 1 plasmid under control of a pLac promoter. Thus, when IPTG is added to the system MS2 is produced and activates the cascade. (B) Expression of GFP and MS2-RFP versus IPTG concentration. As IPTG is increased MS2-RFP is produced which activates the repressor cascade causing GFP expression.

5.3.4 A simplified circuit propagates signal through a transient protein

We next built a simplified circuit to test our ability to propagate a signal through a transient protein. The overall sensing-decoding circuit design is a combination of the two modular circuit architectures encoded in the same cell. The design allows signal propagation from a theophylline input, through the sensing module architecture into MS2 expression, which is then sensed by the decoding circuit and converted into a GFP or RFP output. We constructed

and characterized a simplified version of this overall circuit architecture to confirm that we are able to achieve the desired signal passing between modules (Figure 5.6A).

To construct the minimal sensing-decoding circuit, we combined a single repression sensing circuit and single repression decoding circuit into a two-plasmid system (Figure 5.6A). A high copy plasmid expressed two antisense RNAs: the theophylline activated antisense and the MS2 activated antisense. A low copy plasmid expressed both targets and their outputs: the theophylline antisense represses a dual control repressor regulating MS2-RFP and the MS2 antisense represses a dual control repressor regulating GFP. When theophylline is added, the theophylline antisense will be activated, causing repression of MS2-RFP. This will cause the MS2 activated antisense to become inactive and GFP will be expressed.

We characterized this minimal sensing-decoding circuit by performing a time course experiment on E. coli cultures that contained the circuit plasmids described above. As expected, when theophylline was introduced, we observed MS2-RFP decreasing in expression and increased GFP expression. Thus, we are able to propagate a protein signal between RNA modules. However, there was significant variability amongst the three colonies measured and background fluorescence.

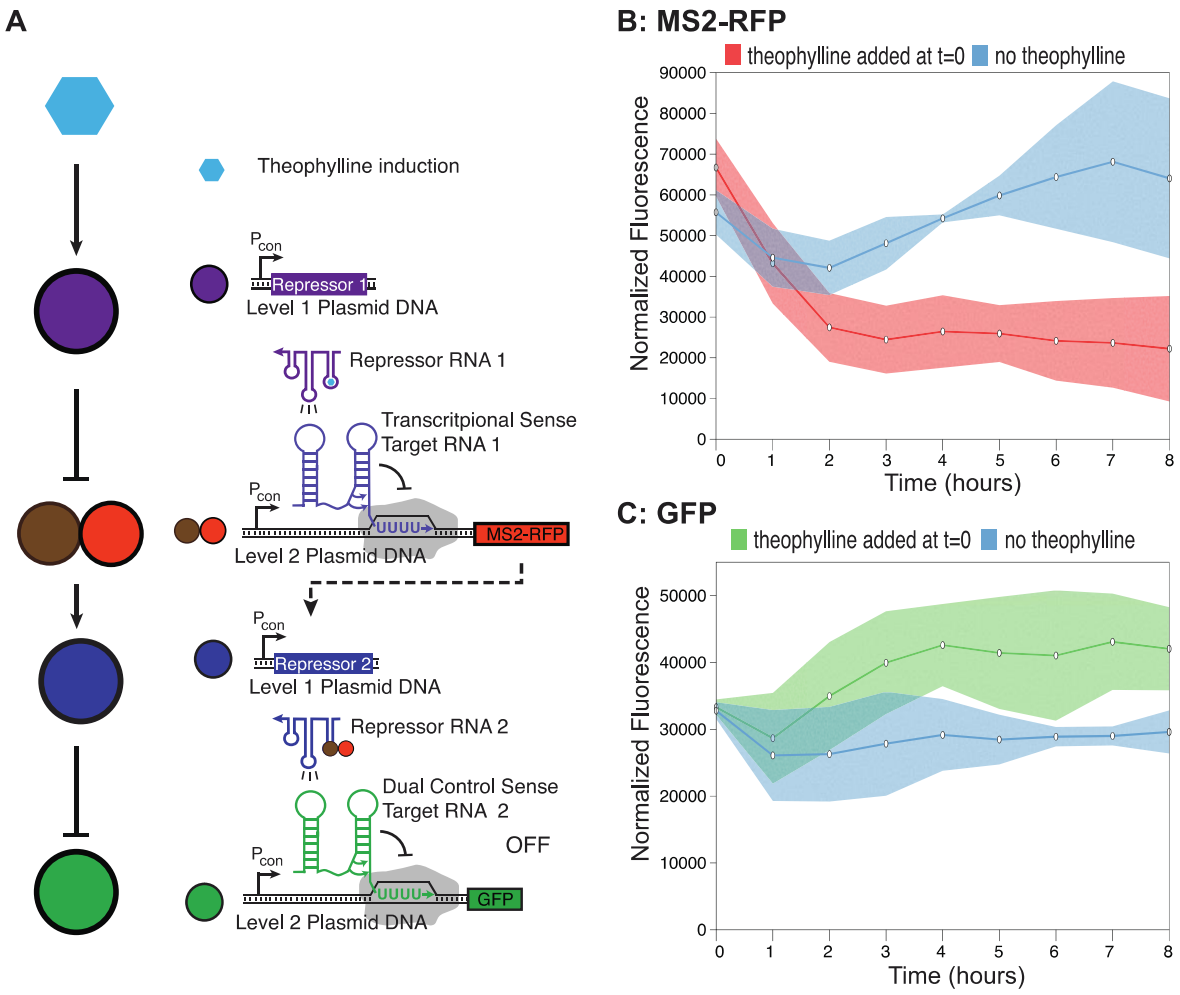


Figure 5.6 A minimal sensing-decoding circuit transmits theophylline signals through MS2 into measurable GFP outputs. (A) Schematic of a simplified temporal signal decoding circuit. This circuit is made up of two RNA repressors buffered by a protein intermediary (MS2). The output is GFP, which is controlled by RNA repressor 2. In the OFF state (no theophylline) RNA repressor 1 is inactive so MS2-RFP is expressed. MS2-RFP activates RNA repressor 2, which represses GFP. When theophylline is added (ON) RNA repressor 1 is activated, repressing MS2-RFP. This causes RNA repressor 2 to become inactive, leading to GFP expression. (B) MS2-RFP fluorescence after induction of the circuit with theophylline. The background (no induction) is shown in blue and is somewhat noisy. MS2-RFP decreases as

expected with induction of theophylline. (C) GFP activates after induction with theophylline. The background (no induction) is shown in blue.

5.3.5 Using a computational model to inform circuit designs.

In addition to studying the proposed circuits experimentally, we have developed a computational framework using a system of kinetic differential equations to qualitatively inform experimental and circuit design. The model uses simplified parameters that capture the overall synthesis and degradation rates of RNAs and proteins but coarse grains biochemical details (Alon 2007). The framework informs experimental design choices such as the optimal timing and length of theophylline pulses required to achieve an output that matches the desired patterns. It also enables the exploration alternative circuit designs without expending resources to build them in cells.

A limitation of the first temporal circuit design is its dependence on the dilution rate of MS2 due to cell division to maintain state and remember the number of pulses. Experiments indicate that the production rate of the fluorescent protein is too similar to the dilution/degradation rate to effectively act as genetic memory (Figure 5.7). While there are ways to modify the dilution rate by slowing growth, an alternative strategy is to design a circuit that remembers state and resets with a separate signal rather than over time. The second design discussed in section 5.3.1 describes such a circuit. Here we present results of modeling this circuit and its internal component, the toggle switch.

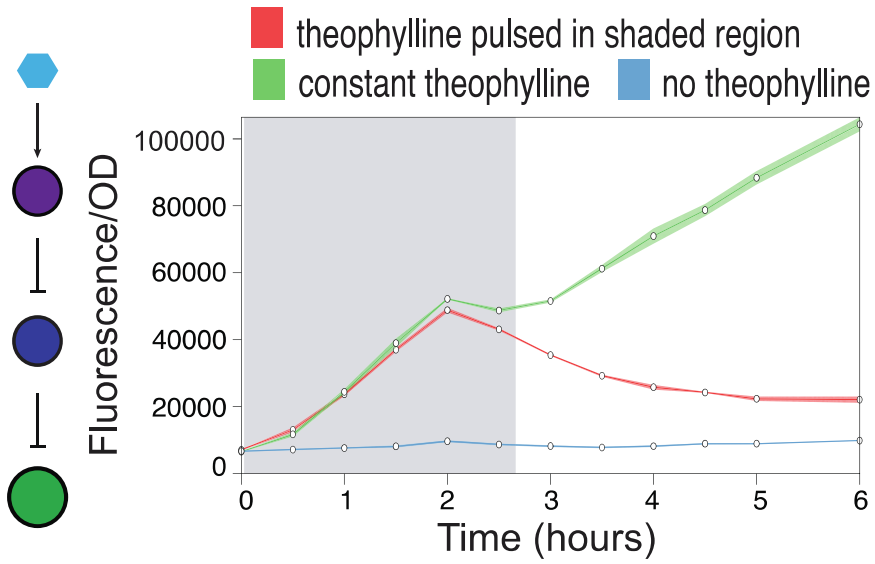


Figure 5.7 A time course using the theophylline cascade indicates that the GFP production rate is similar to its degradation rate. The slope of the theophylline condition (red) from t=0 to 2 hours is similar to the slope after theophylline has been removed (red) from t=2 to 3.5 hours. Experiment performed as in Figure 5.4B.

We began by modeling an RNA toggle switch (Figure 5.8A) controlled by light activation using 6 ODEs (equations 1-6). Light intensity (I_g or green light, I_r for red light) determines the production rate of [A] and [B] (equations 1-2) (Olson & Tabor 2014) based on a published model for the CcaS/CcaR and Cph8/OmpR systems. See tables 1 and 2 in Appendix 1 for variable descriptions and parameter values. Parameters are order of magnitude estimations based on previous studies.

$$\frac{d[A]}{dt} = k_p(I_g)[c_g(t - \tau_{delay}) - A(t)] \quad (1)(\text{Olson et al. 2014})$$

$$\frac{d[B]}{dt} = k_p(I_r)[c_r(t - \tau_{delay}) - B(t)] \quad (2)(\text{Olson et al. 2014})$$

Previous work on modeling RNA interactions indicated that the Hill function accurately describes this system (Hu et al. 2015), so species [C] and [D] are modeled with their interactions according to the network diagram shown in Figure 5.8B. GFP mRNA (M_g) is produced from the same transcript as [D] and shares the same equation.

$$\frac{d[C]}{dt} = \beta\left[1 - \frac{[A]}{K + [A]}\right]^{\eta AC}\left[1 - \frac{[D]}{K + [D]}\right]^{\eta DC} - d_m[C] \quad (3)$$

$$\frac{d[D]}{dt} = \beta\left[1 - \frac{[B]}{K + [B]}\right]^{\eta BD}\left[1 - \frac{[C]}{K + [C]}\right]^{\eta CD} - d_m[D] \quad (4)$$

$$\frac{d[M_g]}{dt} = \beta\left[1 - \frac{[B]}{K + [B]}\right]^{\eta BD}\left[1 - \frac{[C]}{K + [C]}\right]^{\eta CD} - d_m[M] \quad (5)$$

GFP (P_g) production is governed by a translation rate (k_{P_g}) and a degradation rate (d_{P_g}).

$$\frac{d[P_g]}{dt} = k_{P_g}[M_g] - d_{P_g}[P_g] \quad (6)$$

Toggle switches are known to require cooperativity in order to exhibit bistability(Gardner et al. 2000), or the existence of multiple stable states. The pT181 attenuator is not cooperative, but pT181 attenuators used in tandem exhibit cooperative behavior (Lucks et al. 2011) because multiplying non-cooperative hill functions can sharpen transfer curves. By performing a stability analysis around $\frac{d[C]}{dt}$ and $\frac{d[D]}{dt}$ from the toggle switch model, we found conditions for bistability as a function of the number of tandem attenuators for repressors D and C such that $\eta CD = 2$ and $\eta DC = 2$, but $\eta AC = 1$ and $\eta BD = 1$ (Figure 5.9). As shown below, the double attenuator toggle switch exhibits bistability (Figure 5.9B) while the single attenuator has only one stable point (Figure 5.9A). The design is sensitive to changes in K, the repression constant, such that if the repression is poor more tandem attenuators are required to maintain bistability (Figure 5.9C). Additionally, the design is sensitive to changes in β, the transcription rate, such that if the transcription is low bistability is lost (Figure 5.9C). With these tandem attenuators the model indicates that the

proposed network would work as expected such that GFP is expressed after a pulse of green light and repressed after a pulse of red light (Figure 5.8C).

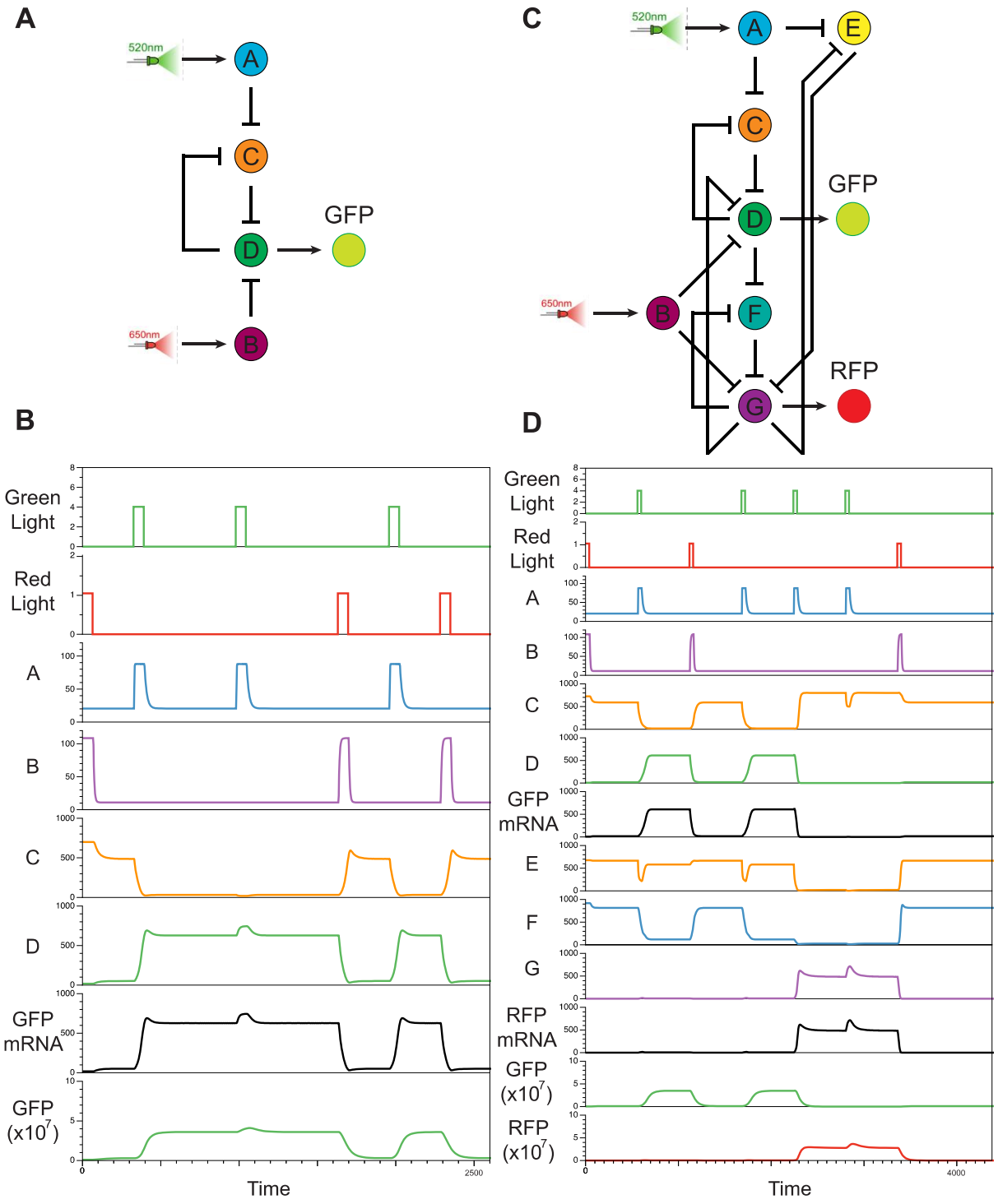


Figure 5.8 Model for the toggle switch and two pulse counter. (A) Network diagram for the toggle switch consisting of two connected repressors responding to green light and red light as inputs. (B) Trajectories for the toggle switch responding to pulses of light. GFP is expressed after a pulse of green light and repressed after a pulse of red light. (C) Network diagram for the two pulse counter made of two connected toggle switches responding to green light as a set and red light as a reset. (D) Trajectories for the two pulse counter responding to pulses of light. GFP is expressed after one pulse of green light, RFP is expressed after two pulses of green light, and red light represses both GFP and RFP.

We then modeled an RNA two pulse counter (Figure 5.8D) controlled by light activation (equations 1-2, 7-15). As with the toggle switch, we used the published light model and constant light intensity to model [A] and [B] production. For this reason, the first two equations modeling the two pulse counter are the same as the toggle switch (equations 1-2). Hill functions were also used to model species [C], [D], [E], [F], [G] and their interactions according to the network diagram shown in Figure 8C. GFP mRNA (M_g) is produced from the same transcript as [D] and RFP mRNA (M_r) is produced from the same transcript as [G].

$$\frac{d[C]}{dt} = \beta\left[1-\frac{[A]}{K+[A]}\right]^{\eta AC}\left[1-\frac{[D]}{K+[D]}\right]^{\eta DC} - d_m[C] \quad (7)$$

$$\frac{d[D]}{dt} = \beta\left[1-\frac{[B]}{K+[B]}\right]^{\eta BD}\left[1-\frac{[C]}{K+[C]}\right]^{\eta CD}\left[1-\frac{[G]}{K+[G]}\right]^{\eta GD} - d_m[D] \quad (8)$$

$$\frac{d[M_g]}{dt} = \beta\left[1-\frac{[B]}{K+[B]}\right]^{\eta BD}\left[1-\frac{[C]}{K+[C]}\right]^{\eta CD}\left[1-\frac{[G]}{K+[G]}\right]^{\eta GD} - d_m[M_g] \quad (9)$$

$$\frac{d[E]}{dt} = \beta\left[1-\frac{[A]}{K+[A]}\right]^{\eta AE}\left[1-\frac{[G]}{K+[G]}\right]^{\eta GE} - d_m[E] \quad (10)$$

$$\frac{d[F]}{dt} = \beta\left[1-\frac{[D]}{K+[D]}\right]^{\eta DF}\left[1-\frac{[G]}{K+[G]}\right]^{\eta GF} - d_m[F] \quad (11)$$

$$\frac{d[G]}{dt} = \beta\left[1-\frac{[B]}{K+[B]}\right]^{\eta BG}\left[1-\frac{[E]}{K+[E]}\right]^{\eta EG}\left[1-\frac{[F]}{K+[F]}\right]^{\eta FG} - d_m[G] \quad (12)$$

$$\frac{d[M_r]}{dt} = \beta\left[1-\frac{[B]}{K+[B]}\right]^{\eta BG}\left[1-\frac{[E]}{K+[E]}\right]^{\eta EG}\left[1-\frac{[F]}{K+[F]}\right]^{\eta FG} - d_m[M_r] \quad (13)$$

GFP and RFP protein (P_g, P_r) production rate is governed by a translation rate (k_{P_g}, k_{P_r}) and a degradation rate (d_{P_g}, d_{P_r}).

$$\frac{d[P_g]}{dt} = k_{P_g}[M_g] - d_{P_g}[P_g] \quad (14)$$

$$\frac{d[P_r]}{dt} = k_{P_r}[M_r] - d_{P_r}[P_r] \quad (15)$$

As in the case of the toggle switch, the two pulse counter requires cooperativity to function properly. In addition to tandem C and D attenuators we found that tandem attenuators are needed for repressors A, G, and E such that the cooperativity constants $\eta CD, \eta DC, \eta AE, \eta GE, \eta GF, \eta EG = 2$, but $\eta AC, \eta BD, \eta GD, \eta DF, \eta BG, \eta FG = 1$. This design is sensitive to changes in K and β similar to the toggle switch. With these tandem repressors the model shows the desired behavior indicating that the proposed network would work as expected: GFP is expressed after one pulse of green light, RFP is expressed after a second pulse of green light, and both GFP and RFP are repressed after a pulse of red light (Figure 5.8D).

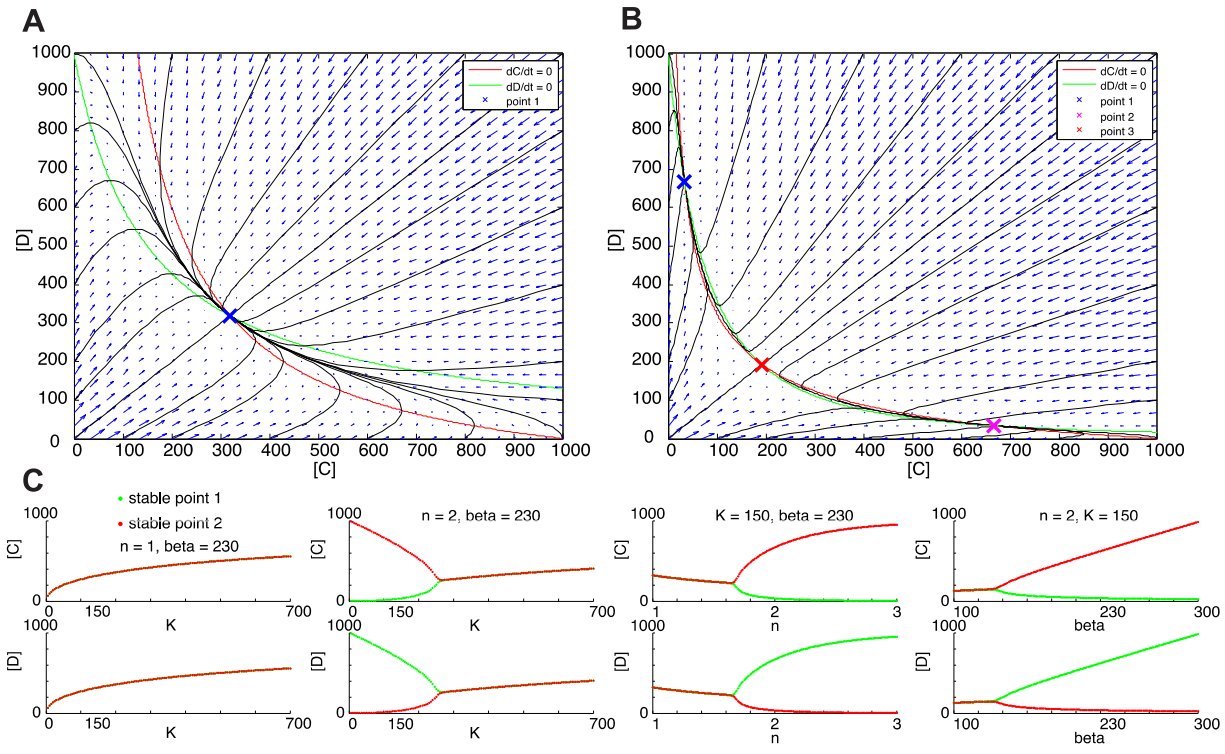


Figure 5.9 Stability analysis of the toggle switch for single attenuators (A) and double attenuators (B) and bifurcations for various values of n, K, and β (C). The nullclines are shown in red (dC/dt=0) and green (dD/dt=0). Solutions with various initial conditions are shown in black and the vector field is shown with blue arrows. The single attenuator toggle switch has one stable point (blue x) while the double attenuator toggle switch has three, two of which are stable (blue x, magenta x) and one which is a saddle point (red x). (C) These bifurcation plots show how the critical points change with change K, n, and β. There is no bistability at n=1 for any value of K. At n=2 bistability appears as K decreases. Increasing values of n and β cause bistability to appear.

5.4 Discussions and Conclusions

In this work we have demonstrated the need for better external communication with cellular systems. We have proposed two possible designs to be constructed with RNA regulators and have outlined a plan to connect light-induction to our RNA circuits to facilitate pulse experiments and reduce the negative metabolic effects of chemical inducers.

Our results have indicated that chemical inducers are non-ideal due to (1) potential toxicity and interference with metabolic processes and (2) the experimental difficulty of pulsing chemical inducers in an appropriate timeframe. Optogenetics offers a unique alternative that could fix both of these issues. Our new designs suggest the CcaS/CcaR and Cph8/OmpR light-switchable systems because they: (i) are compatible with one another giving us two control points, (ii) are light-switchable rather than light-responsive, and (iii) have been successfully used to characterize biological parts. Our designs use a light-switchable, two-component system to drive the activation of RNA circuitry and could later

be used to characterize the temporal dynamics of RNA regulatory parts. This will give us much more flexibility in temporal manipulation and part characterization.

The circuitry developed in this proposal will greatly improve the amount of information that can be transferred to cells by taking advantage of the temporal domain with inducer pulses. The temporal signal decoding circuit will spur development of fundamental principles that can be applied broadly to many challenges in synthetic biology. The RNA circuitry components refined here are modular and can be used for many other applications. This research contributes to a growing discipline of genetic design that will allow scientists to predictably engineer genetic regulatory networks.

5.5 Materials and Methods

Strains, growth medium, and In Vivo *bulk fluorescence time course experiments*.

All experiments were performed in *E. coli* strain TG1. Plasmid combinations were transformed into chemically competent *E. coli* TG1 cells, plated on Difco LB+Agar plates containing appropriate antibiotics and incubated overnight at 37 °C. Transformation plates containing *E. coli* TG1 cells transformed with circuit plasmids were taken out of the incubator and left at room temperature for approximately 3 h. Three colonies were picked and used to inoculate 300 μL of LB containing selective antibiotics in a 2 mL 96-well block (Costar 3960), and grown approximately 17 h overnight. Twenty microliters of each overnight culture was then added to separate wells on a new block containing 980 μL (1:50 dilution) of supplemented M9 minimal media (as mentioned above) containing the selective antibiotics and grown for 4 h at the same conditions as the overnight culture. The optical density (OD, 600 nm) was then measured by transferring 50 μL of culture from the block into a 96-well

plate (Costar 3631) containing 50 µL of phosphate buffered saline (PBS) and measuring using a Biotek Synergy H1m plate reader. The cultures were diluted into 1ml of fresh M9 minimal media to an optical density of 0.015 and grown for four hours. Then theophylline was added to the theophylline condition to a final concentration of 2mM. Every 30 min for the next 4 h, 50 µL from each of the fresh cultures was removed from the 96-well block and transferred to a 96-well plate (Costar 3631) containing 50 µL of phosphate buffered saline (PBS). SFGFP fluorescence (FL, 485 nm excitation, 520 nm emission) and optical density (OD, 600 nm) were then measured at each time point using a Biotek Synergy H1m plate reader.

Bulk fluorescence data analysis.

On each 96-well block, there were two sets of controls; a media blank (M9 alone) and *E. coli* TG1 cells that do not produce SFGFP (transformed with control plasmids JBL001, JBL002, and JBL1856). The block contained three replicates of each control. OD and FL values for each colony at each time point were first corrected by subtracting the corresponding values of the media blank at that same time point. The ratio of FL to OD (FL/OD) was then calculated for each well (grown from a single colony), and the mean FL/OD of TG1 cells without SFGFP at the same time point was subtracted from each colony's FL/OD value to correct for cellular autofluorescence.

5.6 Funding

This work was supported by the Office of Naval Research Young Investigators Program Award (ONR YIP) [N00014-13- 1-0531 to J.B.L.] an NSF CAREER Award [1452441 to J.B.L], and Searle Funds at The Chicago Community Trust [to J.B.L].

5.7 References

Alon, U., 2007. Network motifs: theory and experimental approaches. Nature reviews. Genetics, 8(6), pp.450–461.

Basu, S. et al., 2005. A synthetic multicellular system for programmed pattern formation. Nature, 434(7037), pp.1130–1134.

Baudrimont, A. & Becskei, A., 2015. Gene regulation: Expression feels two pulses. Nature, 527(7576), pp.46–47.

Chappell, J. et al., 2013. The centrality of RNA for engineering gene expression. Biotechnology journal, 8(12), pp.1379–1395.

Friedland, A.E. et al., 2009. Synthetic gene networks that count. Science (New York, N.Y.), 324(5931), pp.1199–1202.

Gardner, T.S., Cantor, C.R. & Collins, J.J., 2000. Construction of a genetic toggle switch in Escherichia coli. Nature, 403(6767), pp.339–342.

Hsiao, V. et al., 2016. A population-based temporal logic gate for timing and recording chemical events. Molecular Systems Biology, 12(5), p.869.

Hu, C.Y., Varner, J.D. & Lucks, J.B., 2015. Generating Effective Models and Parameters for RNA Genetic Circuits. ACS synthetic biology, 4(8), pp.914–926.

Kotula, J.W. et al., 2014. Programmable bacteria detect and record an environmental signal in the mammalian gut. Proceedings of the National Academy of Sciences of the United States of America, 111(13), pp.4838–4843.

Levine, M. & Davidson, E.H., 2005. Gene regulatory networks for development. Proceedings of the National Academy of Sciences of the United States of America, 102(14), pp.4936–4942.

Levskaya, A. et al., 2005. Synthetic biology: engineering Escherichia coli to see light. Nature, 438(7067), pp.441–442.

Lucks, J.B. et al., 2011. Versatile RNA-sensing transcriptional regulators for engineering genetic networks. Proceedings of the National Academy of Sciences of the United States of America, 108(21), pp.8617–8622.

Mutalik, V.K. et al., 2012. Rationally designed families of orthogonal RNA regulators of translation. Nature chemical biology, 8(5), pp.447–454.

Olson, E.J. & Tabor, J.J., 2014. Optogenetic characterization methods overcome key challenges in synthetic and systems biology. Nature chemical biology, 10(7), pp.502–511.

Olson, E.J. et al., 2014. Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. Nature methods, 11(4), pp.449–455.

Rockwell, N.C., Su, Y.-S. & Lagarias, J.C., 2006. Phytochrome structure and signaling mechanisms. Annual review of plant biology, 57(1), pp.837–858.

Schmidl, S.R. et al., 2014. Refactoring and optimization of light-switchable Escherichia coli two-component systems. ACS synthetic biology, 3(11), pp.820–831.

Tabor, J.J., Levskaya, A. & Voigt, C.A., 2011. Multichromatic control of gene expression in Escherichia coli. Journal of Molecular Biology, 405(2), pp.315–324.

Takahashi, M.K. et al., 2015. Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. ACS synthetic biology, 4(5), pp.503–515.

Voigt, C.A., 2006. Genetic parts to program bacteria. Current opinion in biotechnology, 17(5), pp.548–557.

Zoltowski, B.D., Vaccaro, B. & Crane, B.R., 2009. Mechanism-based tuning of a LOV domain

photoreceptor. Nature chemical biology, 5(11), pp.827–834.

CHAPTER 6

TOWARDS AN RNA-BASED BIOLOGICAL CONTROLLER

6.1 Abstract

Controllers are implemented in many engineered systems as a way to control a process regardless of changing environmental conditions. However, controllers have yet to be implemented in a biological context. Here, we show our efforts to construct a biological controller in a cell free transcription-translation (TXTL) system. We propose two controller designs that uses transcriptional riboregulators to compute error and perform proportional adjustment. Previous research has indicated that RNA sequestration must dominate over RNA degradation for such a controller to function. Current repressor sequestration is more efficient than activator sequestration. Due to this, we have found that our repression based controller is able to better track reference point than our activation based controller.

6.2 Introduction

Cells naturally sense and respond to their environment in order to maintain constant internal conditions. Synthetic biologists are starting to explore the idea of synthetic control networks to reduce variability in engineered systems (Dunlop et al. 2010; Ma et al. 2009; Briat, Gupta, et al. 2016; Briat, Zechner, et al. 2016). Reducing uncertainty in engineered systems could vastly improve networks that rely on stable plasmid copy numbers(Park et al. 2002), enzyme concentrations (Dunlop et al. 2010), and regulator concentrations (Del Vecchio et al. 2017). Up to this point, all efforts to build biological controllers have been theoretical, relying on models to determine if a certain controller design is experimentally viable and will produce an effective controller. One such study looked at the ability RNA

regulators, CRISPRi (Qi et al. 2013) and small transcription activating RNAs (STARs) (Chappell et al. 2015), to provide feedback control (Agrawal et al. 2018). This method relies on RNA sequestration to perform error calculation and notably found that it's two designs, a 'direct' controller and an 'indirect' controller could adapt to perturbations in DNA concentration and kinetic parameters. However, one key finding indicated that the RNA sequestration rate must dominate over the RNA degradation rate for effective controller performance.

In industrial systems feedback controllers are used maintain a robust output in the face of environmental uncertainties. Common examples include the speed of your car under auto pilot and the temperature of your house with a thermostat. A user-defined reference point is used to maintain the output of a process at a specified value. The controller calculates the error between the reference and the output using sensors and then performs a correction to the process in order to minimize the error (Figure 6. 1). Ideally the controller should exhibit perfect adaptation, meaning it should be able to track changes in the reference value robustly and show zero steady state error when parameters are perturbed. RNA is an ideal molecule for implementing such control networks in biological systems because of its fast and tunable dynamics (Takahashi et al. 2015; Carrier & Keasling 1999). RNA regulators have recently been developed with dynamic range that challenges protein regulators (Chappell et al. 2017; Green et al. 2014) and are also easier to computationally design.

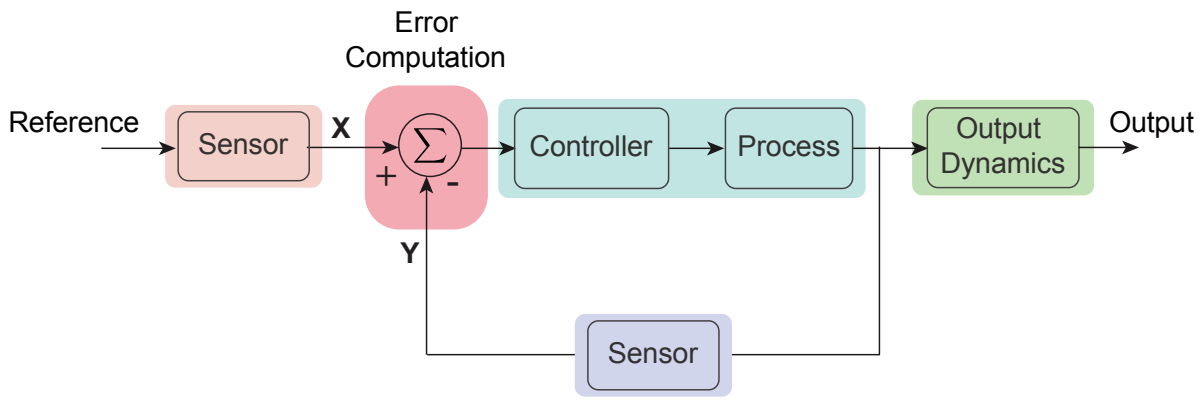


Figure 6.1 Block diagram representation of a closed loop controller from Deepak et al. (Agrawal et al. 2018) In order to maintain a constant output even in changing external conditions, a closed loop is introduced between the process and a reference. The controller calculates an error between the reference and the output and makes a correction based on this error.

Here, we present preliminary work towards implementation of an RNA-based controller. We explore two controller designs: an activation loop controller and a repression loop controller (Figure 6.2). All experiments are performed in TXTL because of its unique ability to allow rapid prototyping of RNA parts and circuits (Takahashi et al. 2015). Each controller calculates error using RNA sequestration. We begin by exploring the activation loop controller that is directly based off of the 'direct' controller modeled in Agrawal et al. (Agrawal et al. 2018) The reference, X, for this controller is a small transcription activating RNA (STAR) that activates expression of Y, that represents both the signal and the output. Y is sequestered by X performing a molecular error calculation and allowing for excess X to correct for external perturbations. Preliminary results for the activation loop controller indicate that STAR sequestration does not dominate over RNA degradation and so the

activation loop design does not function well as a controller. We tried to optimize STAR sequestration extra sequence and promote STAR-sequester binding, but this only resulted in minimal improvements.

The next design we tried was the repression loop controller. Repressor sequestration has already been demonstrated in Chapter 5 and is more efficient than the STAR sequestration developed here. We thought if we could use repressor sequestration instead, this might lead to a more function controller. The repression loop controller uses a repressor sequester as the reference, X, and the repressor as the signal from the output (Y). We found that the repression loop controller shows better reference tracking than the open loop but more work must be done to confirm if the repression loop design is functioning as a robust controller.

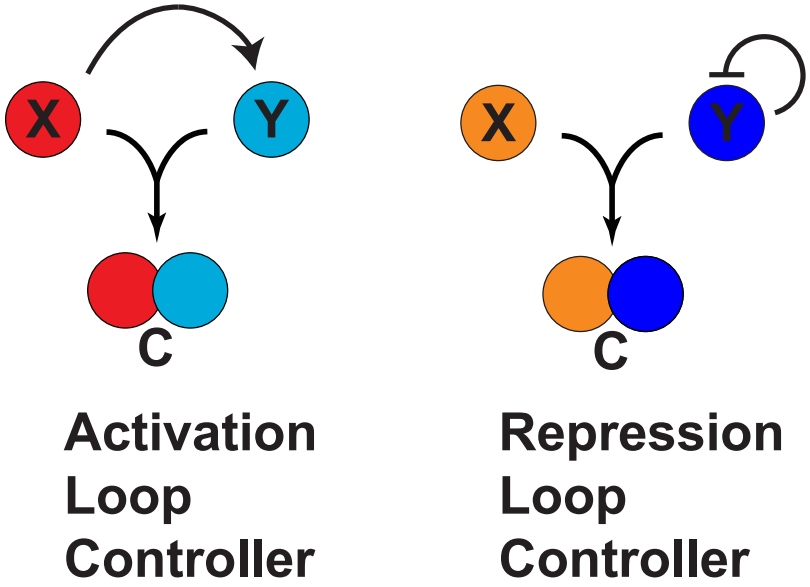


Figure 6.2 Two designs for an RNA-based controller. The activation loop controller (left) uses a STAR while the repression loop controller (right) uses a pT181 repressor.

6.3 Results

6.3.1 Small transcription activating RNA (STAR) sequestration

In order to construct the activation loop controller we need an RNA that can sequester STARs. RNA sequestration has been successfully used to de-repress CRISPRi (Lee et al. 2016) and STAR sequestration was even recently published for the pAD1 STARs (Lee et al. 2018). We began developing STAR sequestration before the pAD1 STAR sequesters were published and choose to build off of the new higher activation, computationally designed stars (Chappell et al. 2017). We designed the sequesters to be complementary to the the linear region of the STAR so they would form a double stranded RNA complex and block STAR from binding its target.

We began by testing STAR sequestration in TXTL. We added plasmids expressing STAR RNA at 8nM, sequester RNA from 4nM to 16nM, and GFP at 1nM to pre-incubated TXTL extract and measured GFP fluorescence every 5 minutes for 3 hours. With no STAR or sequester present the GFP fluorescence was undetectable above background. When STAR was added we see a high ON level of GFP expression. We then added 4nM, 8nM, and 16nM of sequester plasmid and saw 37%, 66%, and 90% repression respectively (Figure 6.3). Ninety percent repression at a 2:1 ratio of sequester to STAR is better than what has been previously published (Lee et al. 2018), but there is still a significant amount of leak given the OFF state of the STAR is essentially zero.

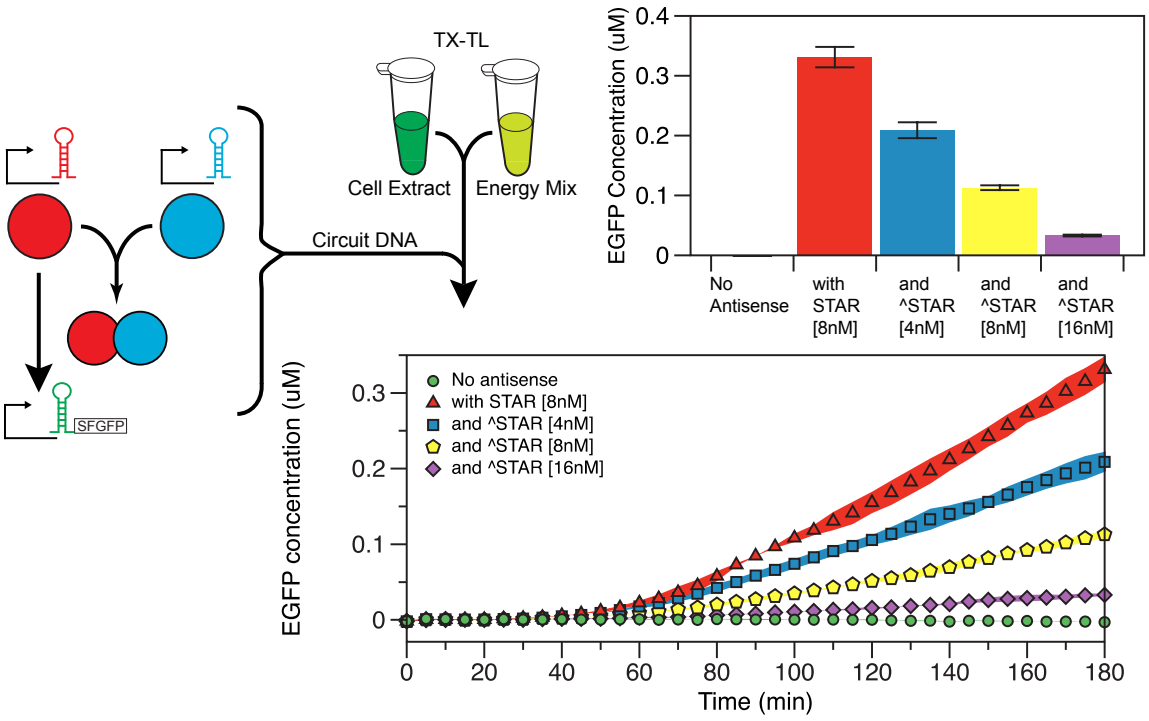


Figure 6.3 STAR sequestration in TXTL. Plasmids expressing the STAR RNA, sequester RNA, and GFP are added to TXTL and then GFP fluorescence is measured every 5 minutes for 3 hours. The STAR sequester binds to the STAR to block activation of GFP. As STAR sequester plasmid concentration increases STAR activation is reduced.

6.3.2 Activation loop controller

The activation loop controller is based off of the 'direct' controller characterized in Agrawal et al. (Agrawal et al. 2018) The goal for this controller is to be able to accurately track changes in the reference and adapt to external perturbations. The controller consists of three plasmids each expressing RNAs that interact create the control circuit. The reference, X, is the DNA expressing a STAR. This STAR activates expression of Y, the DNA expressing the sequester and GFP, the output. A difference between X and Y (error) is detected through

STAR sequestration and a corrective action is taken by the excess or reduced STAR pool (Figure 6.4, left). The open loop controller simply removes the sequestration error calculation mechanism by using a nonfunctional RNA in place of the sequester.

We sought to test the performance of the activation loop controller by changing the reference DNA concentration in TXTL to see if the controller could track the reference and how it performed in comparison to the open loop. We performed a time course experiment in TXTL with 3 difference concentrations of reference: 0nM, 4nM, and 8nM (Figure 6.4, right). We found that the open and closed loops behaved very similarly and both were able to track the reference. At higher concentrations of reference we would expect to see the open loop hit saturation while the closed loop should still be able to track the reference. In order to see this, more experiments would need to be done. However, we can see from this experiment that the controller is not working as expected. If the sequestration was functioning correctly, the closed loop should have lower expression than the open loop and this is now what we see. In order to address this we must improve the sequestration by several orders of magnitude.

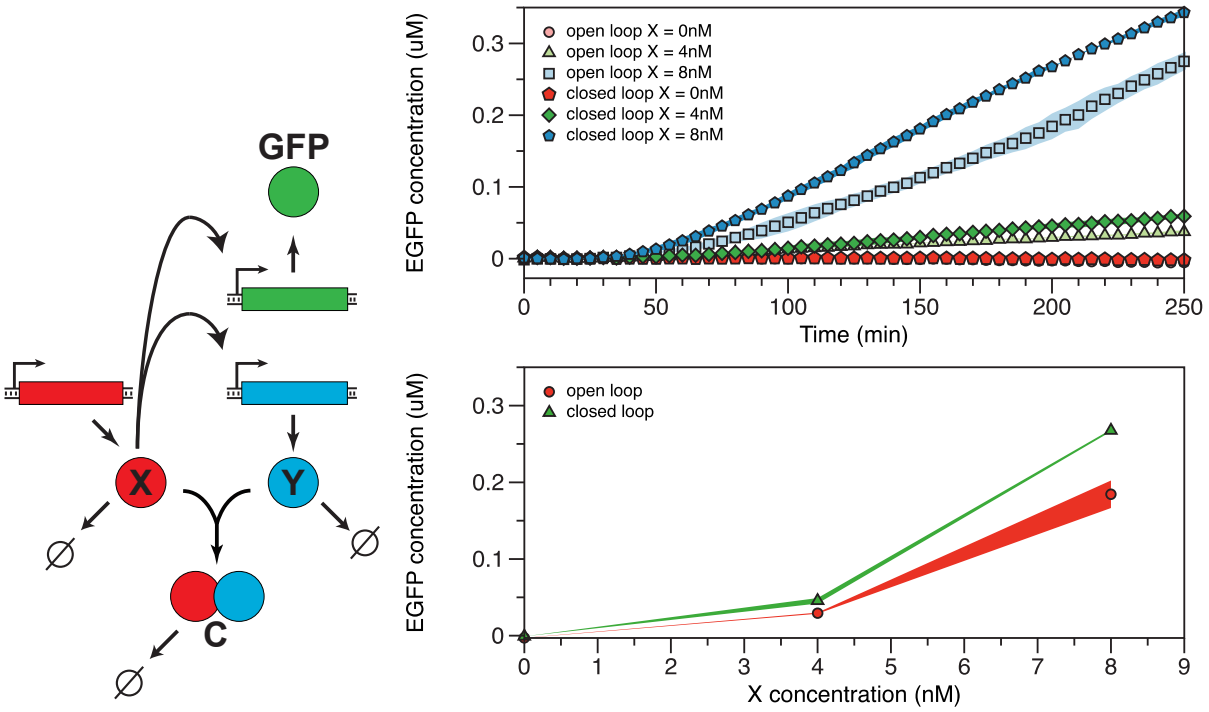


Figure 6.4 The activation loop controller shows poor reference tracking. Schematic for the activation loop controller (left). The activation loop controller consists of three plasmids. The first plasmid expresses STAR RNA (X) that activates the other two plasmids: GFP and the sequester (Y). The sequester (Y) binds to and inhibits the STAR (X). The open loop removes the sequestration between X and Y and instead of a sequester a junk RNA is expressed as Y. At different reference levels, the open and closed loops behave similarly (top right) and show similar levels of referencing tracking (bottom right).

6.3.3 Improving STAR sequestration

Simulations of the activation loop controller indicated that fast sequestration was vital for controller performance (Agrawal et al. 2018). Even at a 2:1 ratio of sequester plasmid to STAR plasmid, the repression is only 90% which is significant leak over the OFF level. Other attempts to sequester STARs have seen similar low efficiencies (Lee et al. 2018). Our

hypothesis for why STAR sequestration works so poorly relies on abortive transcription. When there is no STAR present in the system, RNA Polymerase is still constantly binding to the DNA and creating short abortive transcripts that code for the STAR terminator. This terminator shares sequence with the STAR sequester meaning that when STAR is not present at saturating levels, a STAR sequester-like molecule is being constantly transcribed by RNA polymerase. This will create a threshold that the STAR would need to overcome in order for activation to work and could partially explain why STAR plasmid is needed in great excess to Target-GFP plasmid in order for STAR activation to function. In order for the STAR sequester to have any impact, the STAR sequester plasmid would need to be in even greater excess to create a larger threshold for STAR activation.

We hoped to alleviate this problem by making the binding of STAR to STAR sequester more favorable than STAR to aborted terminator. We used NUPACK, a computation tool for designing RNA structures, to design linear RNA sequences that would not interfere with STAR folding (Zadeh et al. 2011). We then added the complementary sequence to the STAR sequester and tested the new STAR/sequester variants in TXTL (Figure 6.5). At a 1:1 ratio of STAR to sequester the wild type had 3 fold repression while the best variant had 6 fold repression (Figure 6.5). While this is an improvement over the original STAR/sequester pair, we would need several orders of magnitude improvement for the controller to function as expected.

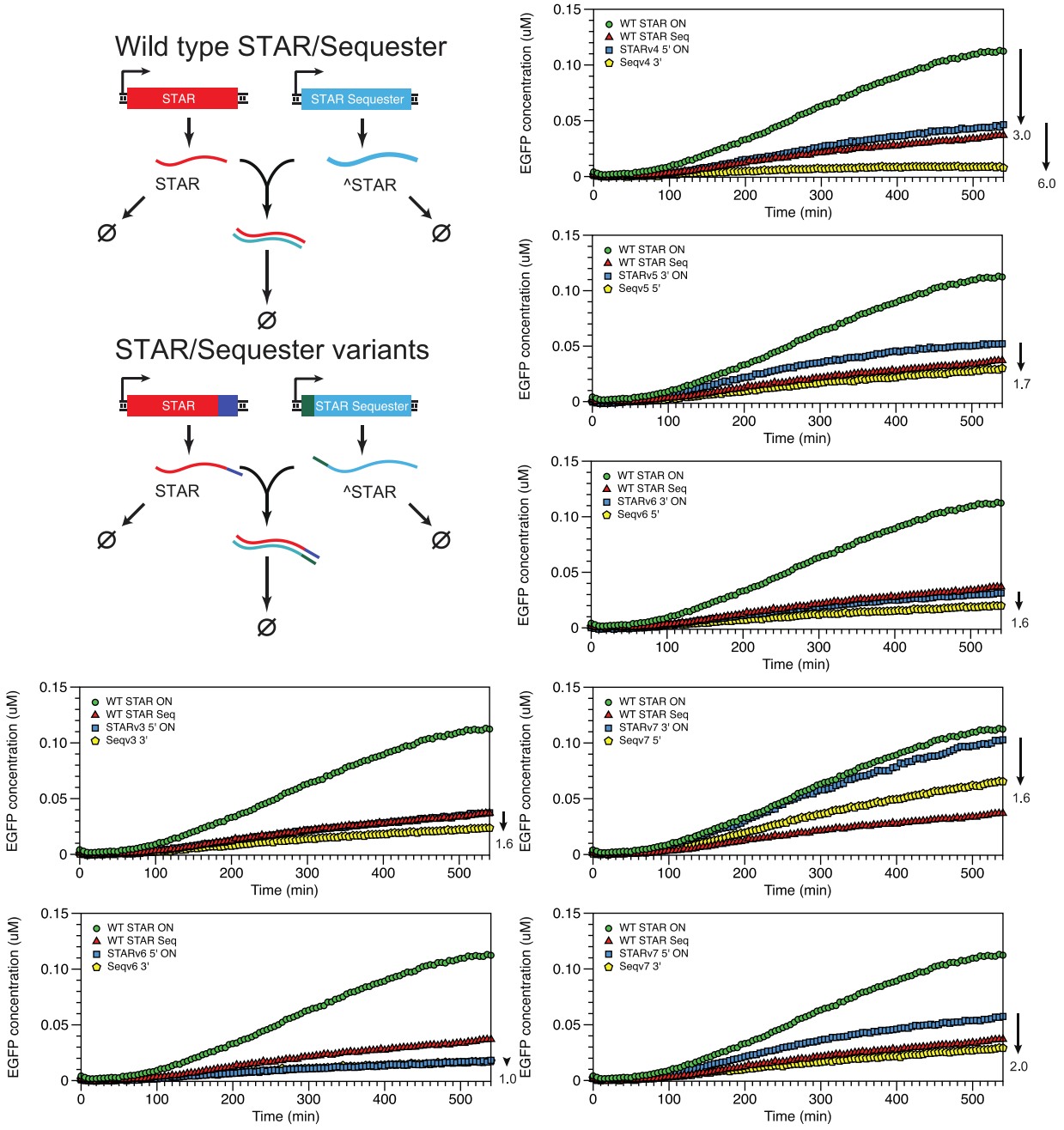


Figure 6.5 New STAR variants do not significantly improve STAR sequestration. We designed new STAR and sequester variants with increased, randomized linear regions to encourage binding between the STAR and the sequester. The wild type STAR sequester system shows 3 fold repression at equimolar concentrations of STAR and repressor plasmid. The best new star/sequester variant shows 6 fold repression.

6.3.4 Repression loop controller

Due to the difficulty of improving STAR sequestration and data suggesting that repressor sequestration is more effective (see chapter 3) we sought to determine if the repression loop controller would function better. This controller combines negative autoregulation and repressor sequestration. Negative autoregulation based on the pT181 attenuator was recently shown to reduce network response time and steady state noise in gene expression (Hu et al. 2018). Like the activation loop controller, this controller also consists of three plasmids that express RNAs to conduct signals and calculate error. The reference is the DNA expressing X, the repressor sequester. The repressor sequester binds to Y, the repressor to calculate error. Y represses itself using negative autoregulation as well as GFP, the output (Figure 6.6, left). The open loop removes the negative feedback on Y. Sequestration is still functional, but the repressor no longer represses itself.

We tested this controller by changing the concentration of reference plasmid to determine if the controller could accurately track the reference in TXTL (Figure 6.6, right) . We found that both the open and closed loops could track changes in the reference, but the closed loop had a more linear response which is what we would expect of a functioning controller. More experiments remain to be done to determine if the controller can respond to environmental perturbations, but the preliminary data is promising, showing an expected difference between the open and closed loops.

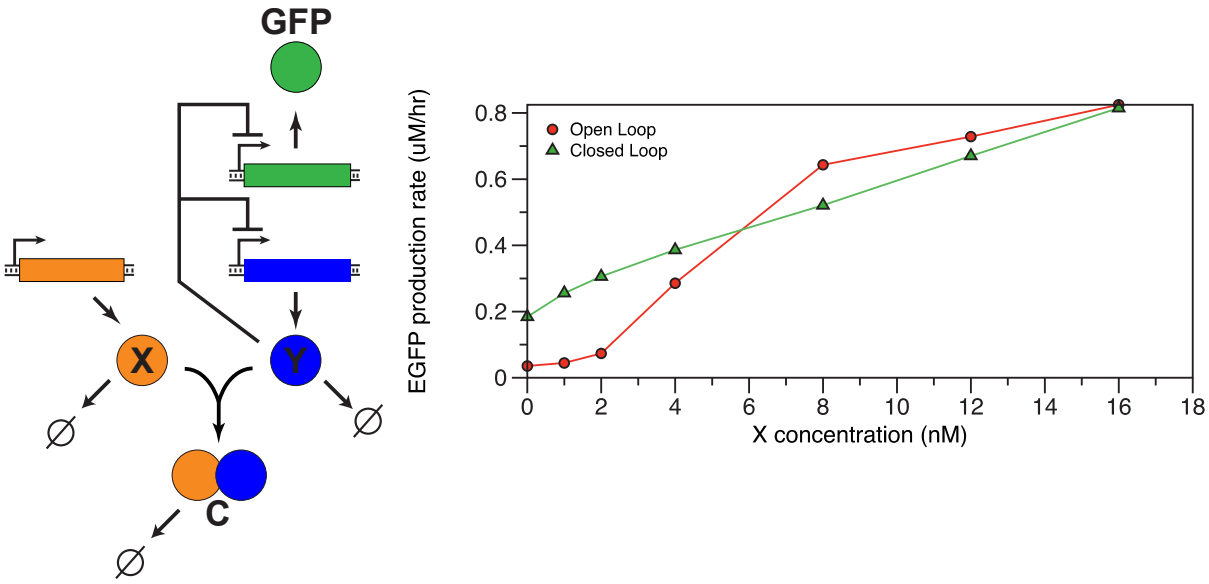


Figure 6.6 The repression loop controller shows more linear reference tracking as compared to the open loop. The repression loop controller consists of three plasmids. The first plasmid expresses the repressor sequester (X) that sequesters the repressor (Y). The repressor (Y) acts as the signal from the output and represses itself as well as a plasmid expressing GFP. The open loop removes the feedback such that the repressor (Y) no longer represses itself. As the concentration of reference (X) increases, the production rate of GFP increases more linearly for the closed loop indicated better controller performance.

6.4 Discussions and Conclusions

In this work we have proposed designs for RNA-based biological controllers, built the necessary components, and presented preliminary data. While the activation-loop controller doesn't show the characteristics we would expect of a functioning controller due to insufficient sequestration, we have shown promising results for the repression loop controller. More work remains to be done to validate the robustness of the repression loop controller. Attempts to improve STAR sequestration to fix the activation loop controller did

not succeed. This marks one of the first steps towards implementing a biological controller in cells and could lead to more predictable and stable genetic networks as networks increase in complexity.

6.5 Materials and Methods

TXTL Extract and Buffer Preparation.

Cell extract and reaction buffer were prepared according to Garamella et al. (Garamella et al. 2016)

TXTL experiments.

TX-TL buffer and extract tubes were thawed on ice for approximately 20 min. Separate reaction tubes were prepared with combinations of DNA representing a given circuit condition. Appropriate volumes of DNA, buffer, and extract were calculated using a custom spreadsheet developed by Sun et al. (Sun et al. 2013) and modified to fit the experiment. Buffer and extract were mixed together and then added to each tube of DNA according to the previously published protocol. Each TX-TL reaction mixture (10 μL each) was transferred to a 384-well plate (Nunc 142761), covered with a plate seal (Nunc 232701), and placed on a Biotek SynergyH1m plate reader. We note that special care is needed when pipetting to avoid air bubbles, which can interfere with fluorescence measurements. Temperature was controlled at 29°C. GFP fluorescence was measured (485 nm excitation, 520 emission) every 5 min.

6.6 Acknowledgements

We thank everyone on the DARPA Biological Control team: Elisa Franco, Mary Dunlop, Xun Tang, Deepak Agrawal, James Chappell, Vincent Noireaux, Chase Beisel, and Colin Maxwel. Special thanks goes to Ryan Marshal for performing the experiment presented in Figure 6.6.

6.7 Funding

This work was supported by the Defense Advanced Research Projects Agency (Contract HR0011-16-C01-34).

6.8 References

Agrawal, D.K. et al., 2018. Mathematical Modeling of RNA-Based Architectures for Closed Loop Control of Gene Expression. ACS synthetic biology, p.acssynbio.8b00040.

Briat, C., Gupta, A. & Khammash, M., 2016. Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Biomolecular Networks. Cell Systems, 2(1), pp.15–26.

Briat, C., Zechner, C. & Khammash, M., 2016. Design of a Synthetic Integral Feedback Circuit: Dynamic Analysis and DNA Implementation. ACS synthetic biology, 5(10), pp.1108–1116.

Carrier, T.A. & Keasling, J.D., 1999. Library of synthetic 5' secondary structures to manipulate mRNA stability in Escherichia coli. Biotechnology progress, 15(1), pp.58–64.

Chappell, J. et al., 2017. Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. Nature Communications, 8(1), p.795.

Chappell, J., Takahashi, M.K. & Lucks, J.B., 2015. Creating small transcription activating RNAs. Nature chemical biology, 11(3), pp.214–220.

Del Vecchio, D. et al., 2017. A Blueprint for a Synthetic Genetic Feedback Controller to Reprogram Cell Fate. Cell Systems, 4(1), pp.109–120.e11.

Dunlop, M.J., Keasling, J.D. & Mukhopadhyay, A., 2010. A model for improving microbial biofuel production using a synthetic feedback loop. Systems and synthetic biology, 4(2), pp.95–104.

Garamella, J. et al., 2016. The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. ACS synthetic biology, 5(4), pp.344–355.

Green, A.A. et al., 2014. Toehold switches: de-novo-designed regulators of gene expression. Cell, 159(4), pp.925–939.

Hu, C.Y. et al., 2018. Engineering a Functional small RNA Negative Autoregulation Network with Model-guided Design. ACS synthetic biology, p.acssynbio.7b00440.

Lee, Y.J. et al., 2016. Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. Nucleic acids research, 44(5), pp.2462–2473.

Lee, Y.J., Kim, S.-J. & Moon, T.S., 2018. Multilevel Regulation of Bacterial Gene Expression with the Combined STAR and Antisense RNA System. ACS synthetic biology, 7(3), pp.853–865.

Ma, W. et al., 2009. Defining Network Topologies that Can Achieve Biochemical Adaptation. Cell, 138(4), pp.760–773.

Park, K. et al., 2002. Origin pairing ("handcuffing") as a mode of negative control of P1 plasmid copy number. The EMBO journal, 21(3), pp.481–481.

Qi, L.S. et al., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152(5), pp.1173–1183.

Sun, Z.Z. et al., 2013. Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. Journal of visualized experiments : JoVE, (79), p.e50762.

Takahashi, M.K. et al., 2015. Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. ACS synthetic biology, 4(5), pp.503–515.

Zadeh, J.N. et al., 2011. NUPACK: Analysis and design of nucleic acid systems. Journal of computational chemistry, 32(1), pp.170–173.

CHAPTER 7

CONCLUSIONS

In this work we have optimized RNA regulators, built new ways to tune them, and used them to create synthetic networks with novel behaviors. We have improved the dynamic range of the pT181 transcriptional attenuator by adding translational regulation and using it in a RNA-only network to reduce circuit leak. The dual control strategy has also been applied to STARs and a library of orthogonal attenuators. In order to tune our regulators we developed a method of sequestration that allows us to shift the regulator transfer function. This method is broadly applicable and can also be used on orthogonal attenuators without inhibiting their orthogonality. The modularity of RNA regulators was demonstrated in our pulse generating network that uses the timescale difference between CRISPRi and STARs to create a pulse of gene expression. The modeling tools developed here can be used to combine other regulators as networks grow in complexity.

Additionally, we explore emerging network concepts including how to pass information to cells given the limited numbers of inducers and biological control. In order to send multiple signals to a cell using one inducer we explore the idea of encoding information in the temporal domain with pulses of a single inducer. We develop two new circuit designs and weigh their advantages and disadvantages. Models of control networks taught us that that sequestration is a valid way to perform error calculation in a controller, but sequestration strength is important for controller functionality. Though more work needs to be done, we perform experiments that represent some of the first attempts to build controllers in the lab. This research marks a significant step towards engineering sophisticated RNA networks with predictable behaviors.

APPENDIX A

Supplementary Information for Achieving large dynamic range control of gene expression with a compact RNA transcription-translation regulator

A.1 Supplementary Tables, Figures, and Notes

	Description
Table A1	Important DNA sequences
Table A2	Plasmids used in this study
Figure A1	Plasmid architectures used in this study
Note A1	Supplementary Materials and Methods
Figure A2	Day 1, 2, and 3 of repressor cascade
Figure A3	Repressor cascade from glycerol stock
Figure A4	RepC Knockout
Figure A5	qRT-PCR
Figure A6	Transcriptional activator mechanism
Figure A7	In vivo expression data used to calculate orthogonality matrices
Figure A8	Sequence and structure of the pT181 dual control attenuator
Figure A9	Sequence and structure of the pT181 antisense and truncated pT181 antisense
Figure A10	Induction curves
Table A3	Averages and standard deviations for Figure 2.2A and 2.2B
Figure A11	Sequence and structure of the pT181 dual control activator
Table A4	Averages and standard deviations for Figure 2.3B and 2.3C
Figure A12	Sequence and structure of interaction regions for antisense mutants and fusions
Table A5	Averages and standard deviations for Figure 2.4B and 2.4C
Table A6	Standard deviations for orthogonality matrix repression Figure 2.5B
Table A7	Standard deviations for orthogonality matrix repression Figure 2.5C
References	

Table A1: Important DNA sequences

Name	Sequence
J23119	TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT
pLac Promoter	AATTGTGAGCGGATAACAATTGACATTGTGAGCGGATAACAAGATACT
Super folder green fluorescent protein (Ribosome binding site (RBS) - SFGFP)	AGGAGGAAGGATCTATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTC CCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTC CGTGGAGAGGGTGAAGGTGATGCTACAAACGGAAAACTCACCCTTAAATT TATTTGCACTACTGGAAAACTACCTGTTCCGTGGCCAACACTTGTCACTAC TCTGACCTATGGTGTTCAATGCTTTTCCCGTTATCCGGATCACATGAAACG GCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCAC TATATCTTTCAAAGATGACGGGACCTACAAGACGCGTGCTGAAGTCAAGTT TGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAGGGTATTGATTTTA AAGAAGATGGAAACATTCTTGGACACAAACTCGAGTACAACTTTAACTCA CACAATGTATACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAA CTTCAAAATTCGCCACAACGTTGAAGATGGTTCCGTTCAACTAGCAGACCA TTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAA CCATTACCTGTCGACACAATCTGTCCTTTCGAAAGATCCCAACGAAAAGCG TGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGG CATGGATGAGCTCTACAAA
TrrnB	GAAGCTTGGGCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGC CGTCGACCATCATCATCATCATCATTGAGTTTAAACGGTCTCCAGCTTGGC TGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA ACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGT GGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGA TGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAA TAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTT TGTCGGTGAACT
Monomeric Red fluorescent protein (mRFP)	ATGGCAAGTAGCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCG TATGGAAGGTTCCGTTAACGGTCACGAGTTCGAAATCGAAGGTGAAGGTG AAGGTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACCAAAG GTGGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAGTACG GTTCCAAAGCTTACGTTAAACACCCGGCTGACATCCCGGACTACCTGAAAC TGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAACTTCGAAGACG GTGGTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCA TCTACAAAGTTAAACTGCGTGGTACCAACTTCCCGTCCGACGGTCCGGTTA TGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCGAACGTATGTACCCGG AAGACGGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAAGAC GGTGGTCACTACGACGCTGAAGTTAAAACCACCTACATGGCTAAAAAACCG GTTCAGCTGCCGGGTGCTTACAAAACCGACATCAAACTGGACATCACCTCC CACAACGAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAAGGTCGT CACTCCACCGGTGCTTAATAA
LacI ORF	GGCACGTAAGAGGTTCCAACTTTCACCATAATGAAACATACTAGAGAAAG AGGAGAAATACTAGATGGTGAATGTGAAACCAGTAACGTTATACGATGTC GCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAG GCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCG GAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAG TCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGC AAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGG

	TGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGC ACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGG ATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGT TATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCC ATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACC AGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGC GTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGC AAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATC AGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTG GTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTT ATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAA CCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCA ATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCA ATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGG CACGACAGGTTTCCCGACTGGAAAGCGGGCAGGCTGCAAACGACGAAAACT ACGCTTTAGTAGCTTAATAACTCTGATAGTGCTAGTGTAGATCCCTACTAG AGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCG TTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTC ACCTTCGGGTGGGCCTTTCTGCGTTTATA
Theophylline aptamer-pT181-mutant antisense (aptamer-antisense-sTRSV Ribozyme)	GGTGATACCAGCATCGTCTTGATGCCCTTGGCAGCACCTCTTTGAATGGTG CTGCCCTGCAACTTTGGCGAGGGACAGGGCGACTCCTTTTTATTTCTGTCA CCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAG
sTRSV Ribozyme	CTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGAAACAG
Double pT181 antisense (BamHI-BglII Scar-sTRSV Ribozyme antisense)x 2	GGATCTCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGA AACAGGGATCTATACAAGATTATAAAAACAACTCAGTGTTTTTTTCTTTG AATGATGTCGTTCACAAACTTTGGTCAGGGCGTGAGCGACTCCTTTTTATT TGGATCTCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACG AAACAGGGATCCTAACTCGAGATACAAGATTATAAAAACAACTCAGTGTT TTTTTCTTTGAATGATGTCGTTCACAAACTTTGGTCAGGGCGTGAGCGACT CCTTTTTATTTGGATCT
pT181 repressor (sense target – *repC* 96nt fragment)	AACAAAATAAAAAGGAGTCGCTCACGCCCTGACCAAAGTTTGTGAACGAC ATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTATATTTA GATATTAAACGATATTTAAATATACATAAAGATATATATTTGGGTGAGCG ATTCCTTAAACGAAATTGAGATTAAGGAGTCGCTCTTTTTTATGTATAAA AACAATCATGCAAATCATTCAAATCATTTGGAAAATCACGATTTAGACAA TTTTTCTAAAACCGGCTACTCTAATAGCCGGTTGTAAGGATCT
pT181 repressor	AACAAAATAAAAAGGAGTCGCTCACGCCCTGACCAAAGTTTGTGAACGAC ATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTATATTTA

with WT terminator (sense target – *repC* 96nt fragment)	GATATTAAACGATATTTAAATATACATAAAGATATATATTTGGGTGAGCG ATTCCTTAAACGAAATTGAGATTAAGGAGTCGATTTTTTATGTATAAAAA CAATCATGCAAATCATTCAAATCATTTGGAAAATCACGATTTAGACAATT TTTCTAAAACCGGCTACTCTAATAGCCGGTTGTAAGGATCT
pT181-mutant 1 repressor (sense target – *repC* 96nt fragment)	AACAAAATAAAAAGGAGTCGCTCTGTCCCTCGCCAAAGTTGCAGAACGACA TCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTATATTTAG ATATTAAACGATATTTAAATATACATAAAGATATATATTTGGGTGAGCGA TTCCTTAAACGAAATTGAGATTAAGGAGTCGCTCTTTTTTATGTATAAAA ACAATCATGCAAATCATTCAAATCATTTGGAAAATCACGATTTAGACAAT TTTTCTAAAACCGGCTACTCTAATAGCCGGTTGTAAGGATCT
pT181-mutant 2 repressor (sense target – *repC* 96nt fragment)	AACAAAATAAAAAGGAGTCGCTCGTACCCTCTGCAAAGTTAACGAACGAC ATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTATATTTA GATATTAAACGATATTTAAATATACATAAAGATATATATTTGGGTGAGCG ATTCCTTAAACGAAATTGAGATTAAGGAGTCGCTCTTTTTTATGTATAAA AACAATCATGCAAATCATTCAAATCATTTGGAAAATCACGATTTAGACAA TTTTTCTAAAACCGGCTACTCTAATAGCCGGTTGTAAGGATCT
Fusion 3 repressor (sense target – *repC* 96nt fragment)	AACAAAATAAAAAGGAGTCGCTCACGCCTCGAACTTGGCGGAACGCAGTGT GAACGACATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGT ATATTTAGATATTAAACGATATTTAAATATACATAAAGATATATATTTGG GTGAGCGATTCCTTAAACGAAATTGAGATTAAGGAGTCGCTCTTTTTTATG TATAAAAACAATCATGCAAATCATTCAAATCATTTGGAAAATCACGATTT AGACAATTTTTCTAAAACCGGCTACTCTAATAGCCGGTTGTAAGGATCT
Fusion 4 repressor (sense target – *repC* 96nt fragment)	AACAAAATAAAAAGGAGTCGCTCACGTTCAACTTTGGCGAGTACGATGTG AACGACATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTA TATTTAGATATTAAACGATATTTAAATATACATAAAGATATATATTTGGG TGAGCGATTCCTTAAACGAAATTGAGATTAAGGAGTCGCTCTTTTTTATGT ATAAAAACAATCATGCAAATCATTCAAATCATTTGGAAAATCACGATTTA GACAATTTTTCTAAAACCGGCTACTCTAATAGCCGGTTGTAAGGATCT
pT181 activator (sense target – *repC* 96nt fragment)	TTGGGTGAGCGATTCCTTAAACGAAATTGAGATTAAGGAGTCGCTCTTTTT TTTTTATGTATAAAAACAATCATGCAAATCATTCAAATCATTTGGAAAAT CACGATTTAGACAATTTTTCTAAAACCGGCTACTCTAATAGCCGGTTGTAA GGATCT
pT181 dual control activator (sense target – *repC* 12nt fragment)	TTGGGTGAGCGATTCCTTAAACGAAATTGAGATTAAGGAGTCGATTTTTT ATGTATAAAAAC
pT181 dual control repressor with WT	AACAAAATAAAAAGGAGTCGCTCACGCCCTGACCAAAGTTTGTGAACGAC ATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTATATTTA GATATTAAACGATATTTAAATATACATAAAGATATATATTTGGGTGAGCG ATTCCTTAAACGAAATTGAGATTAAGGAGTCGATTTTTTATGTATAAAAA

terminator (sense target – *repC* 12nt fragment)	C
pT181 dual control repressor (sense target – *repC* 12nt fragment)	AACAAAATAAAAAGGAGTCGCTCACGCCCTGACCAAAGTTTGTGAACGAC ATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTATATTTA GATATTAAACGATATTTAAATATACATAAAGATATATATTTGGGTGAGCG ATTCCTTAAACGAAATTGAGATTAAGGAGTCGCTCTTTTTTATGTATAAA AAC
pT181 mutant 1 dual control repressor (sense target – *repC* 12nt fragment)	AACAAAATAAAAAGGAGTCGCTCTGTCCCTCGCCAAAGTTGCAGAACGACA TCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTATATTTAG ATATTAAACGATATTTAAATATACATAAAGATATATATTTGGGTGAGCGA TTCCTTAAACGAAATTGAGATTAAGGAGTCGATTTTTTATGTATAAAAAC
pT181 mutant 2 dual control repressor (sense target – *repC* 12nt fragment)	AACAAAATAAAAAGGAGTCGCTCGTACCCTCTGCAAAGTTAACGAACGAC ATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTATATTTA GATATTAAACGATATTTAAATATACATAAAGATATATATTTGGGTGAGCG ATTCCTTAAACGAAATTGAGATTAAGGAGTCGATTTTTTATGTATAAAAA C
Fusion 3 dual control repressor (sense target – *repC* 12nt fragment)	AACAAAATAAAAAGGAGTCGCTCACGCCTCGAACTTGGCGGAACGCAGTGT GAACGACATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGT ATATTTAGATATTAAACGATATTTAAATATACATAAAGATATATATTTGG GTGAGCGATTCCTTAAACGAAATTGAGATTAAGGAGTCGATTTTTTATGT ATAAAAAC
Fusion 4 dual control repressor (sense target – *repC* 12nt fragment)	AACAAAATAAAAAGGAGTCGCTCACGTTCAACTTTGGCGAGTACGATGTG AACGACATCATTCAAAGAAAAAAACACTGAGTTGTTTTTATAATCTTGTA TATTTAGATATTAAACGATATTTAAATATACATAAAGATATATATTTGGG TGAGCGATTCCTTAAACGAAATTGAGATTAAGGAGTCGATTTTTTATGTA TAAAAAC
pT181 antisense	ATACAAGATTATAAAAACAACTCAGTGTTTTTTTCTTTGAATGATGTCGT TCACAAACTTTGGTCAGGGCGTGAGCGACTCCTTTTTATTTGGATCT
pT181	ATACAAGATTATAAAAACAACTCAGTGTTTTTTTCTTTGAATGATGTCGT

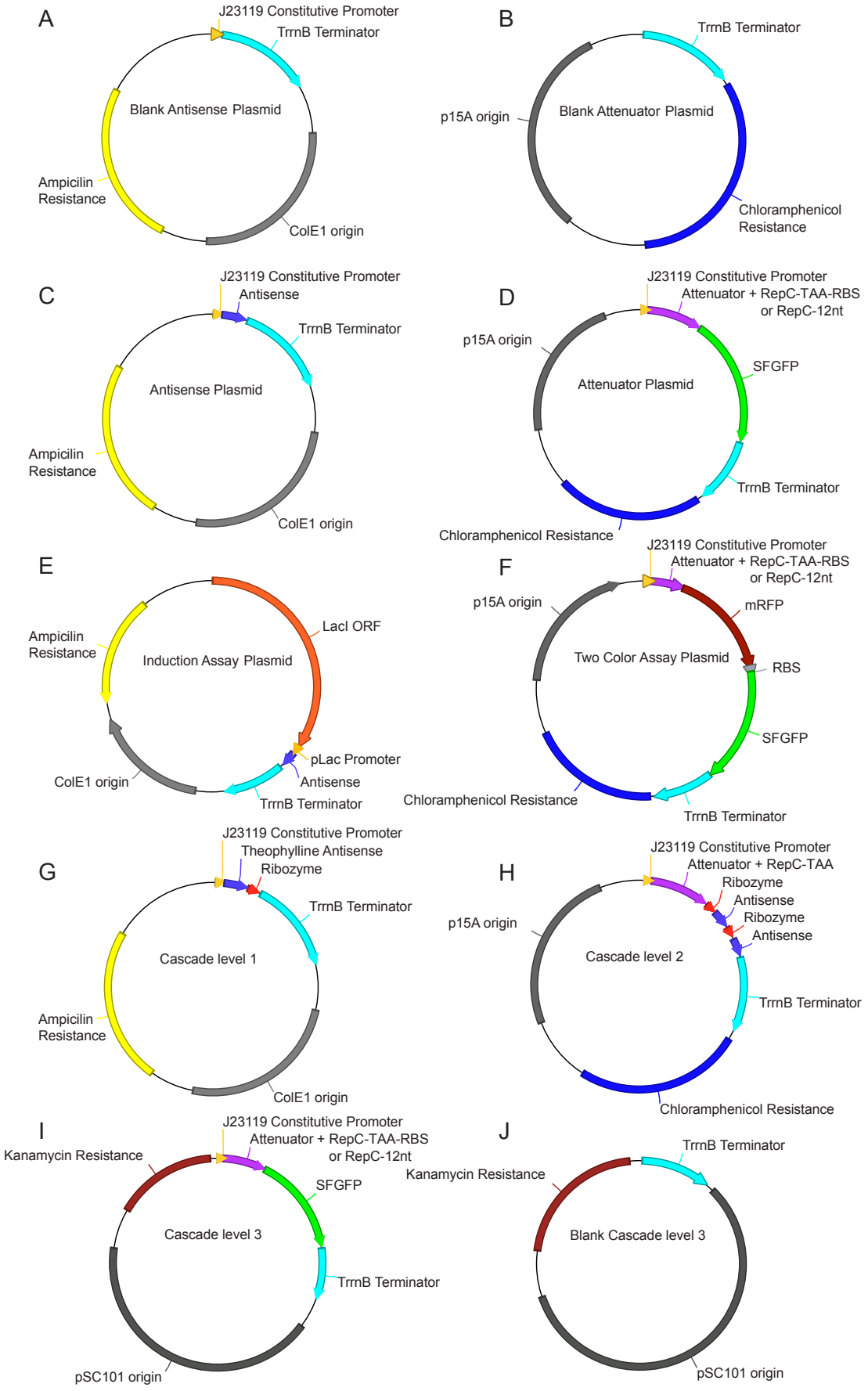
mutant 1 antisense	TCTGCAACTTTGGCGAGGGACAGAGCGACTCCTTTTTATTTGGATCT
pT181 mutant 2 antisense	ATACAAGATTATAAAAACAACTCAGTGTTTTTTTCTTTGAATGATGTCGT TCGTTAACTTTGCAGAGGGTACGAGCGACTCCTTTTTATTTGGATCT
Fusion 4 antisense	ATACAAGATTATAAAAACAACTCAGTGTTTTTTTCTTTGAATGATGTCGT TCACATCGTACTCGCCAAAGTTGAACGTGAGCGACTCCTTTTTATTTGGAT CT
Fusion 3 antisense	ATACAAGATTATAAAAACAACTCAGTGTTTTTTTCTTTGAATGATGTCGT TCACACTGCGTTCCGCCAAGTTCGAGGCGTGAGCGACTCCTTTTTATTTGG ATCT
pT181 activator antisense	AACAAAATAAAGCAATAAGGAATCGCTCACCCAAAGGATCT
pT181 truncated antisense	TGAATGATGTCGTTCACAAACTTTGGTCAGGGCGTGAGCGACTCCTTTTTG GATCT
Fusion 3 truncated antisense	TGAATGATGTCGTTCACACTGCGTTCCGCCAAGTTCGAGGCGTGAGCGACT CCTTTTTGGATCT
pT181 mutant 1 truncated antisense	TGAATGATGTCGTTCTGCAACTTTGGCGAGGGACAGAGCGACTCCTTTTTG GATCT
pT181 mutant 2 truncated antisense	TGAATGATGTCGTTCGTTAACTTTGCAGAGGGTACGAGCGACTCCTTTTTG GATCT
Fusion 4 truncated antisense	TGAATGATGTCGTTCACATCGTACTCGCCAAAGTTGAACGTGAGCGACTCC TTTTTGGATCT

Table S2 – Plasmids used in this study. Sequences in the plasmid architecture can be found in Table A1.

Plasmid #	Plasmid architecture	Name	Figure	Figure S1 map
JBL001	TrrnB – CmR – p15A origin	CmR/p15A control	2-6, S2-5, S7, S10	B
JBL002	J23119 – TrrnB – ColE1 origin – AmpR	No antisense control	2-6, S2-5, S7, S10	A
JBL003	J23119 – pT181 sense target – *repC*(1-96) – RBS – SFGFP – TrrnB – CmR – p15A origin	pT181 repressor WT term	2	D
JBL004	J23119 – pT181 antisense – TrrnB – ColE1 origin – AmpR	pT181 antisense	2, 4, 5, S4, S5, S7	C
JBL006	J23119 – pT181 sense target mutant terminator –	pT181	2, 4, S5,	D

	repC(1-96) – RBS – SFGFP – TrrnB – CmR – p15A origin	repressor mut term	S10	
JBL007	J23119 – pT181 mutant 1 sense target mutant terminator – *repC*(1-96) – RBS – SFGFP – TrrnB – CmR – p15A origin	pT181 mutant 1 repressor	4	D
JBL008	J23119 – pT181 mutant 1 antisense – TrrnB – ColE1 origin – AmpR	pT181 mutant 1 antisense	4, 5, S7	C
JBL009	J23119 – pT181 mutant 2 sense target mutant terminator – *repC*(1-96) – RBS – SFGFP – TrrnB – CmR – p15A origin	pT181 mutant 2 repressor	4	D
JBL010	J23119 – pT181 mutant 2 antisense – TrrnB – ColE1 origin – AmpR	pT181 mutant 2 antisense	4, 5, S7	C
JBL021	LacI ORF – pLac – pT181 antisense – TrrnB – ColE1 origin – AmpR	IPTG inducible antisense	S10	E
JBL1033	J23119 – Fusion 4 antisense – TrrnB – ColE1 origin – AmpR	Fusion 4 antisense	4, 5, S7	C
JBL1035	J23119 – Fusion 3 antisense – TrrnB – ColE1 origin – AmpR	Fusion 3 antisense	4, 5, S7	C
JBL1039	J23119 – Fusion 3 sense target mutant terminator – *repC*(1-96) – RBS – SFGFP – TrrnB – CmR – p15A origin	Fusion 3 repressor	4	D
JBL1126	J23119 – Fusion 4 sense target mutant terminator – *repC*(1-96) – RBS – SFGFP – TrrnB – CmR – p15A origin	Fusion 4 repressor	4	D
JBL1843	J23119 – theophylline aptamer-pT181 mutant 1 antisense – sTRSV ribozyme – TrrnB – ColE1 origin – AmpR	Aptamer pT181 mutant 1 antisense, aptamer-AS-2	6, S2, S3	G
JBL1844	J23119 – pT181 mutant 1 sense target mutant terminator – *repC*(1-96) – (sTRSV ribozyme – pT181 antisense)x2 – TrrnB – CmR – p15A origin	Cascade L2 on p15A/CmR backbone	6, S2, S3	H
JBL1855	J23119 – pT181 sense target – *repC*(1-96) – RBS – SFGFP – TrrnB – pSC101 origin – KanR	pT181 repressor	6, S2, S3	I
JBL1856	TrrnB – pSC101 origin – KanR	pSC101/KanR control	6, S2, S3	J
JBL2071	J23119 – pT181 activator – RBS – SFGFP – TrrnB – CmR – p15A origin	pT181 activator	3	D
JBL2128	J23119 – pT181 activator antisense – TrrnB – ColE1 origin – AmpR	pT181 activator antisense	3	C
JBL2412	J23119 – pT181 dual control sense target – *repC*(1-12) – SFGFP – TrrnB – CmR – p15A origin	pT181 DC repressor WT term	2, 4, 5, S4, S5, S7, S10	D

JBL2413	J23119 – pT181 dual control sense target mutant terminator – *repC*(1-12) – SFGFP – TrrnB – CmR – p15A origin	pT181 DC repressor	2	D
JBL2415	J23119 – pT181 dual control mutant 1 sense target – *repC*(1-12) – SFGFP – TrrnB – CmR – p15A origin	pT181 DC mutant 1 repressor	4, 5, S7	D
JBL2421	J23119 – pT181 dual control activator sense target – *repC*(1-12) – SFGFP – TrrnB – CmR – p15A origin	pT181 DC activator	4, 5, S7	D
JBL2427	J23119 – Fusion 3 dual control sense target – *repC*(1-12) – SFGFP – TrrnB – CmR – p15A origin	Fusion 3 DC repressor	4, 5, S7	D
JBL2432	J23119 – pT181 dual control mutant 2 sense target – *repC*(1-12) – SFGFP – TrrnB – CmR – p15A origin	pT181 DC mutant 2 repressor	4, 5, S7	D
JBL2434	J23119 – Fusion 4 dual control sense target – *repC*(1-12) – SFGFP – TrrnB – CmR – p15A origin	Fusion 4 DC repressor	4, 5, S7	D
JBL2463	J23119 – pT181 dual control sense target – *repC*(1-12) – RFP – RBS – SFGFP – TrrnB – CmR – p15A origin	Dual control RFP/GFP	2	F
JBL2464	J23119 – pT181 truncated antisense – TrrnB – ColE1 origin – AmpR	pT181 truncated antisense	5, S7	C
JBL2469	J23119 – Fusion 3 truncated antisense – TrrnB – ColE1 origin – AmpR	Fusion 3 truncated antisense	5, S7	C
JBL2489	J23119 – pT181 mutant 1 truncated antisense – TrrnB – ColE1 origin – AmpR	pT181 mutant 1 truncated antisense	5, S7	C
JBL2490	J23119 – pT181 mutant 2 truncated antisense – TrrnB – ColE1 origin – AmpR	pT181 mutant 2 truncated antisense	5, S7	C
JBL2491	J23119 – Fusion 4 truncated antisense – TrrnB – ColE1 origin – AmpR	Fusion 4 truncated antisense	5, S7	C
JBL2493	J23119 – pT181 dual control sense target – *repC*(1-12) – SFGFP – TrrnB – pSC101 origin – KanR	pT181 DC WT Term repressor	6, S2, S3	I
JBL2526	J23119 – pT181 dual control sense target – SFGFP – TrrnB – CmR – p15A origin	pT181 DC repressor RepC KO	S4	D


A
J23119 Constitutive Promoter
TrrnB Terminator
Blank Antisense Plasmid
Ampicilin Resistance
ColE1 origin
B
TrrnB Terminator
p15A origin
Blank Attenuator Plasmid
Chloramphenicol Resistance
C
J23119 Constitutive Promoter
Antisense
TrrnB Terminator
Antisense Plasmid
Ampicilin Resistance
ColE1 origin
D
J23119 Constitutive Promoter
Attenuator + RepC-TAA-RBS or RepC-12nt
p15A origin
SFGFP
Attenuator Plasmid
TrrnB Terminator
Chloramphenicol Resistance
E
Ampicilin Resistance
LacI ORF
Induction Assay Plasmid
pLac Promoter
ColE1 origin
Antisense
TrrnB Terminator
F
J23119 Constitutive Promoter
Attenuator + RepC-TAA-RBS or RepC-12nt
p15A origin
mRFP
Two Color Assay Plasmid
RBS
SFGFP
Chloramphenicol Resistance
TrrnB Terminator
G
J23119 Constitutive Promoter
Theophylline Antisense Ribozyme
TrrnB Terminator
Cascade level 1
Ampicilin Resistance
ColE1 origin
H
J23119 Constitutive Promoter
Attenuator + RepC-TAA
Ribozyme
Antisense
Ribozyme
Antisense
p15A origin
Cascade level 2
TrrnB Terminator
Chloramphenicol Resistance
I
J23119 Constitutive Promoter
Attenuator + RepC-TAA-RBS or RepC-12nt
Kanamycin Resistance
SFGFP
Cascade level 3
TrrnB Terminator
pSC101 origin
J
TrrnB Terminator
Kanamycin Resistance
Blank Cascade level 3
pSC101 origin

Supplementary Figure A1: Plasmid architectures for plasmids used in this study. (A) Antisense plasmid blank control (B) Attenuator plasmid blank control (C) Antisense plasmid architecture (D) Attenuator plasmid architecture (E) Induction assay plasmid architecture (F) Two color assay plasmid architecture (G) Architecture of level 1 (top level) of the cascade (H) Architecture of level 2 (mid level) of the cascade (I) Architecture of level 3 (bottom level) of the cascade (J) Cascade level 3 (bottom level) blank control plasmid. Specific sequences can be found in Table A1.

Supplementary Note A1: Supplementary Materials and Methods.

Total RNA extraction for quantitative PCR.

qRT-PCR was performed following the MIQE guidelines (1). Strain, transformation, media, and growth conditions were all the same as for end point experiments described in the Materials in Methods in the main text. Plasmids were transformed, and subsequent colonies were grown overnight as described for in vivo bulk fluorescence measurements. For each biological replica, 20 µl of a single overnight culture was added to three wells containing 980 µl (1:50 dilution) of supplemented M9 minimal medium containing the selective antibiotics and grown for 4 h at the same conditions as the overnight cultures. For each plasmid combination, 500 µl of cells were removed from three wells (grown from one colony) and combined into a 1.6-ml tube and pelleted by centrifugation at 13,000 r.p.m. for 1 min. Total RNA extraction was performed using Trizol reagent (Life Technologies) and an ethanol precipitation as described in Chappell et al. (2). The supernatant was removed, and the remaining pellet was resuspended in 750 µl of Trizol reagent (Life

Technologies), homogenized by repetitive pipetting, and incubated at room temperature for 5 min. 150 µl of chloroform was added, and the samples were mixed for 15 s and incubated at room temperature for 3 min. Following incubation, the samples were centrifuged for 15 min at 12,000g at 4 °C, and 200 µl of the top aqueous layer was removed. One microliter of glycogen (20 µg/µl; Life Technologies) and 375 µl of isopropanol were added to the aqueous phase, and the sample was incubated at room temperature for 10 min and centrifuged for 15 min at 15,000 r.p.m. at 4 °C. Following centrifugation, the isopropanol was carefully removed from the total RNA/glycogen pellets, washed in 600 µl of chilled 70% ethanol (EtOH) and centrifuged for 2 min at 15,000 r.p.m. at 4 °C. EtOH was removed, and tubes were centrifuged for another 2 min at 15,000 r.p.m. at 4 °C to ensure that all of the ethanol was effectively removed. Pellets were resuspended in 20 µl of RNase free double-distilled water (ddH2O) and stored at -80 °C.

DNase treatment of total RNA for qPCR.

The total RNA was treated with Turbo DNase to remove any genomic or plasmid DNA from the RNA extraction (2). Purified total RNA samples were quantified by the Qubit Fluorometer (Life Technologies) and were diluted to a concentration of 30 ng/µl in a total of 10 µl RNase free ddH2O and digested by Turbo DNase (Life Technologies) according to the manufacturer's protocol. After digestion, 150 µl of RNase free ddH2O and 200 µl phenol/chloroform was added, and the sample was vortexed for 10 s and incubated for 3 min at room temperature and centrifuged for 10 min at 15,000 r.p.m. at 4 °C. After centrifugation, 190 µl of the top aqueous layer was carefully removed, 190 µl of chloroform was added, and samples were vortexed for 10 s, incubated for 3 min at room temperature and centrifuged for 10 min at

15,000 r.p.m. at 4 °C. After centrifugation, 170 µl of the top aqueous layer was carefully removed, 170 µl of chloroform was added, and samples were vortexed for 10 s, incubated for 3 min at room temperature and centrifuged for 10 min at 15,000 r.p.m. at 4 °C. After centrifugation, 120 µl of the top aqueous layer was carefully removed and added to 1 µl glycogen, 360 µl of chilled 100% EtOH and 12 µl of 3 M sodium acetate, pH 5. Samples were vortexed for 10 s and stored at −80 °C for 1 h. Samples were then centrifuged for 30 min at 15,000 r.p.m. at 4 °C. Supernatant was removed, and the pellets were washed in 600 µl of chilled 70% EtOH. Samples were then centrifuged for 2 min at 15,000 r.p.m. at 4 °C, and the EtOH was removed. Samples were recentrifuged for 2 min at 15,000 r.p.m. at 4 °C, and residual EtOH was removed, and pellets were air-dried for 10 min, and eluted in 10 µl RNase fee ddH2O and stored at -80 °C.

Normalization of total RNA, reverse transcription and qPCR measurements.

To enable comparison between different samples, each DNase treated sample was normalized to contain the same total RNA concentration. Each sample was quantified by Qubit Fluorometer, and the sample was diluted to 1 ng/µl of total RNA in 12 µl RNase free ddH2O. For the reverse transcription, one microliter of this total RNA, 1 µl of 2 µM reverse transcription primer (RT SFGFP: TTATTTGTAGAGCTCATCCATG), 1 µl of 10 mM of dNTPs (New England BioLabs) and RNase-free ddH2O (up to 6.5 µl) were incubated for 5 min at 65 °C and cooled on ice for 5 min. 0.25 µl of Superscript III reverse transcriptase (Life Technologies), 1 µl of 100 mM Dithiothreitol (DTT), 1 µl first-strand buffer (Life Technologies), 0.5 µl RNaseOUT (Life Technologies) and RNase free H2O up to 3.5 µl were then added, and the solution was incubated at 55 °C for 1 h, 75 °C for 15 min and then stored at

−20 °C. qPCR was performed using 5 µl of Maxima SYBR green qPCR master mix (Thermo Scientific), 1 µl of cDNA and 0.5 µl of 2 µM SFGFP qPCR primers (SFGFP.Fwd: CACTGGAGTTGTCCCAATTCT, SFGFP.Rev: TCCGTTTGTAGCATCACCTTC) and RNase-free ddH2O up to 10 µl. A Bio-Rad CFX Connect Real-Time System (Bio-Rad) was used for data collection using the following PCR program: 50 °C for 2 min, 95 °C for 10 min, followed by 30 cycles of 95 °C for 15 s and 60 °C for 1 min. All of the measurements were followed by melting curve analysis. A Hard-Shell 96-well PCR Plate (HSL9641, Bio-Rad) and a Microseal 'B' seal (MSB1001, Bio-Rad) were used for all measurements. Results were analyzed using Bio-Rad CFX Manager (V 3.1, Bio-Rad) by a relative standard curve. For quantification, a six-point standard curve covering a 100,000-fold range of SFGFP DNA concentrations (R2 > 0.99) was run in parallel and used to determine the relative SFGFP cDNA abundance in each sample. It was shown that the SFGFP qPCR primer set had a primer efficiency between 101–104%. All of the cDNA samples were measured in triplicate, and nontemplate controls run in parallel to control for contamination and nonspecific amplification or primer dimers. All NTC samples were quantified after 30 cycles. In addition, qPCR was performed on total RNA samples to confirm limited plasmid DNA contamination of cDNA samples. Control samples were quantified at least 4 cycles after their cDNA samples. Melting curve analysis was performed to confirm that only a single product was amplified. Units are reported as ng/µl x 10-7 representing cDNA concentration in the processed samples.

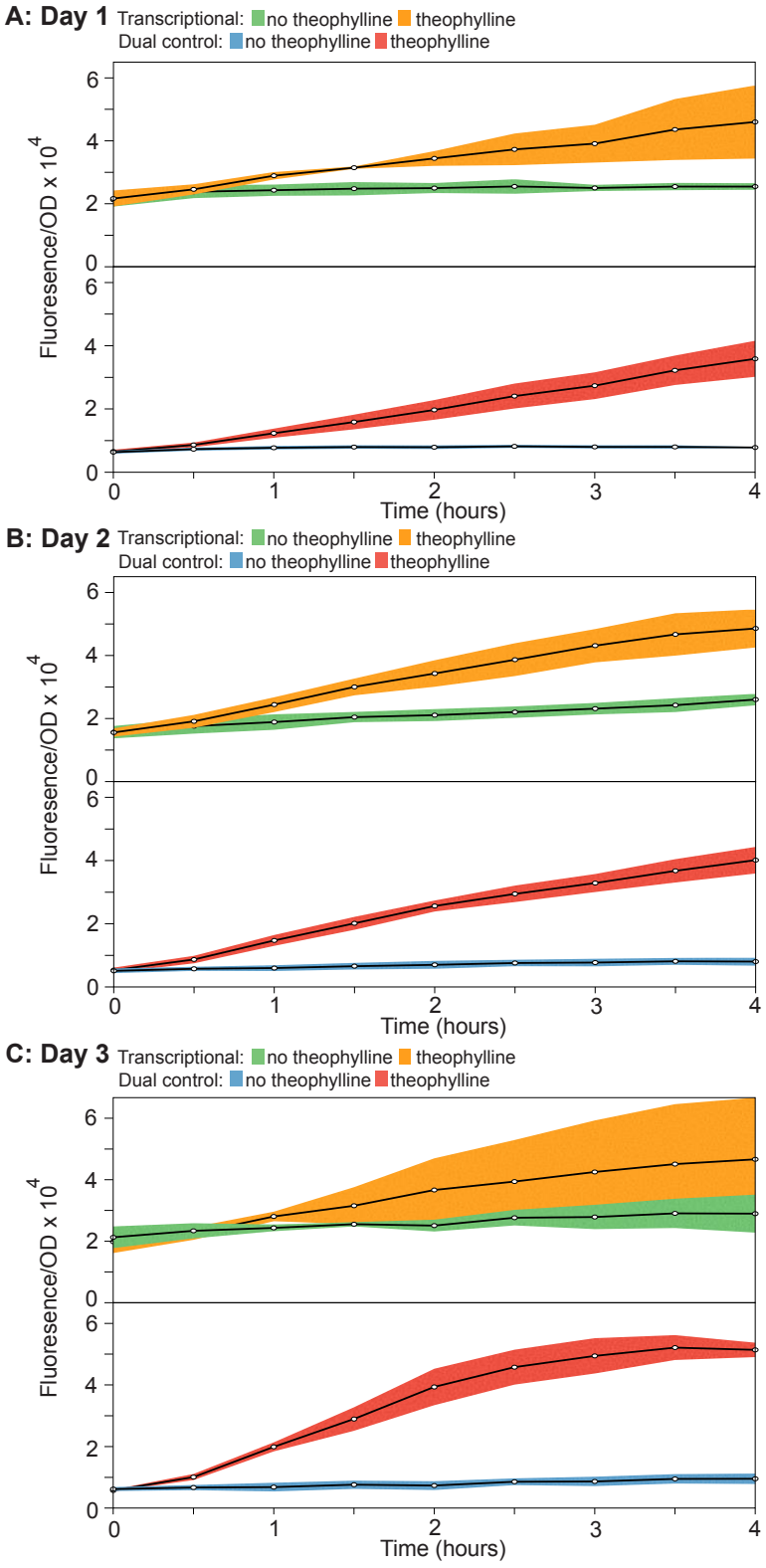
Induction curve assay.

Strain, transformation, media, and growth conditions were all the same as for end point experiments described in the Materials and Methods in the main text. Plasmid combinations were transformed into chemically competent E. coli TG1 cells, plated on Difco LB+Agar plates containing selective antibiotics and incubated overnight (approximately 17 hours) at 37°C. Plates were taken out of the incubator and left at room temperature for approximately 7 h. Three colonies were used to separately inoculate 300 μL of LB containing selective antibiotics in a 2 mL 96-well block (Costar 3960), and grown approximately 17 h overnight at 37°C. Four microliters of this overnight culture was then added to 196 μL (1:50 dilution) of supplemented M9 minimal media containing the selective antibiotics and the required amount of IPTG. After 4 h of incubation at 37C, 50 μL of this culture was then transferred to a 96-well plate (Costar 3631) containing 50 μL of phosphate buffered saline (PBS). Fluorescence (485 nm excitation, 520 nm emission) and optical density (OD, 600 nm) were then measured using a Biotek SynergyH1m plate reader.

RepC Knockout Assay.

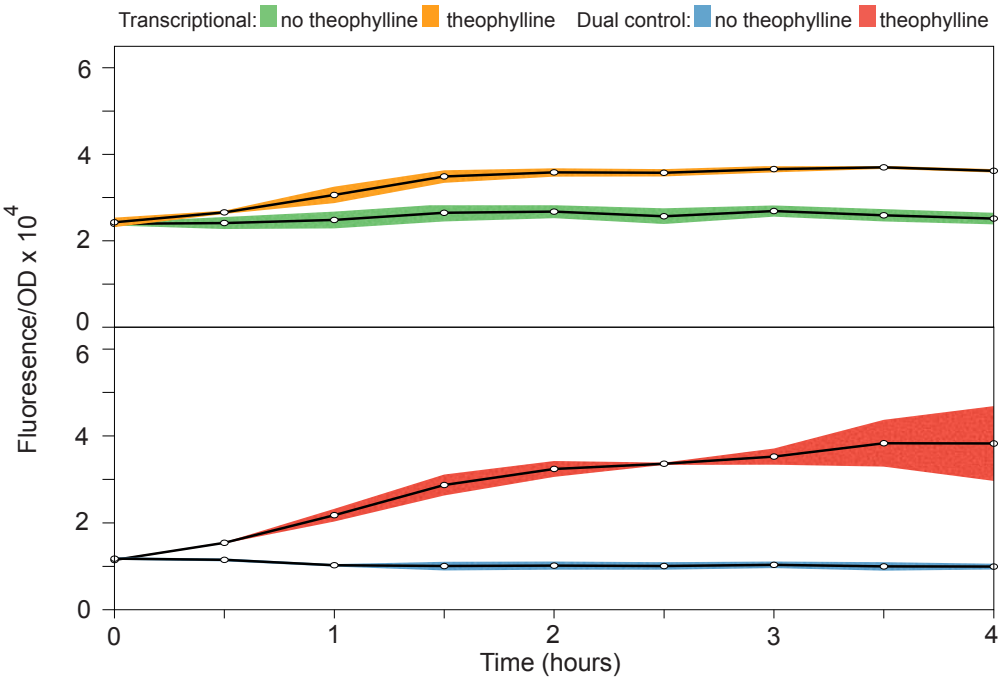
Strain, Transformation, media, and growth conditions were all the same for the end point experiments described in the Materials and Methods in the main text. After the 4 h incubation at 37C in supplemented M9 minimal media, 50 μL of this culture was then transferred to a 96-well plate (Costar 3631) containing 50 μL of phosphate buffered saline (PBS). Fluorescence (485 nm excitation, 520 nm emission) and optical density (OD, 600 nm) were then measured using a Biotek SynergyH1m plate reader.

In Vivo bulk fluorescence time course experiments from glycerol stocks, Strain, transformation, media, and growth conditions were all the same as for end point experiments described in the Materials in Methods in the main text. Transformation plates containing E. coli TG1 cells transformed with three cascade plasmids (Supplementary Table A2) were taken out of the incubator and left at room temperature for approximately 7 h. Colonies were picked and incubated at 37C overnight in LB containing selective antibiotics. The following morning, 500µL of culture was mixed with 500µL of 50% glycerol and frozen at -80C. Three colonies were picked from the frozen glycerol stocks and used to inoculate 300 µL of LB containing selective antibiotics in a 2 mL 96-well block (Costar 3960), and grown approximately 17 h overnight at the same conditions as described for an end point experiment (see Materials and Methods in main text). Cultures were diluted 1:50 in separate wells on a new block containing 1 mL of supplemented M9 minimal media containing the selective antibiotics and grown for 4 h at the same conditions as the overnight culture. Then theophylline was added to the theophylline condition to a final concentration of 2mM. Every 30 min for the next 4 h, 50 µL from each of the fresh cultures was removed from the 96-well block and transferred to a 96-well plate (Costar 3631) containing 50 µL of phosphate buffered saline (PBS). SFGFP fluorescence (FL, 485 nm excitation, 520 nm emission) and optical density (OD, 600 nm) were then measured at each time point using a Biotek Synergy H1m plate reader.

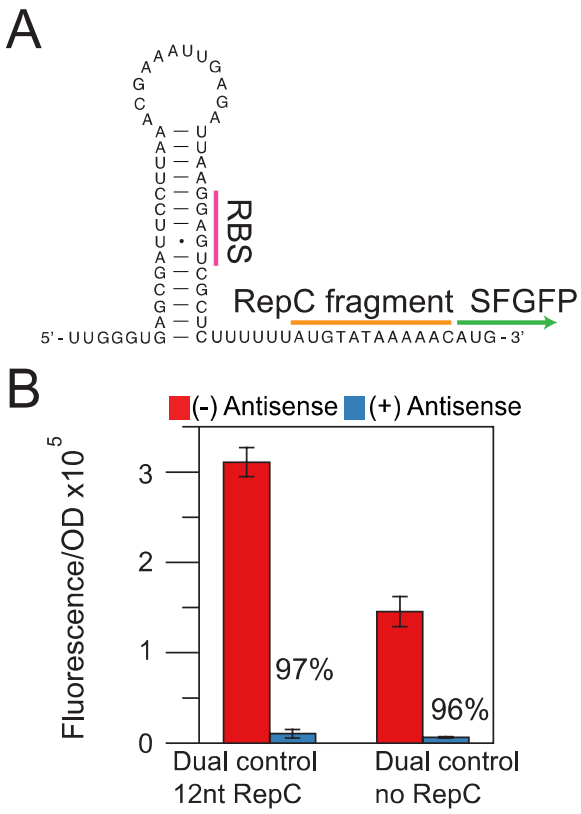


Supplementary Figure A2: Additional data for repressor cascade (Figure 2.6). Functional data for the transcriptional and dual control repressor cascades over time. All three plasmids were co-transformed into *E. coli* TG1 cells. Theophylline (2mM) is spiked at t=0 hours (orange for transcriptional and red for dual control)

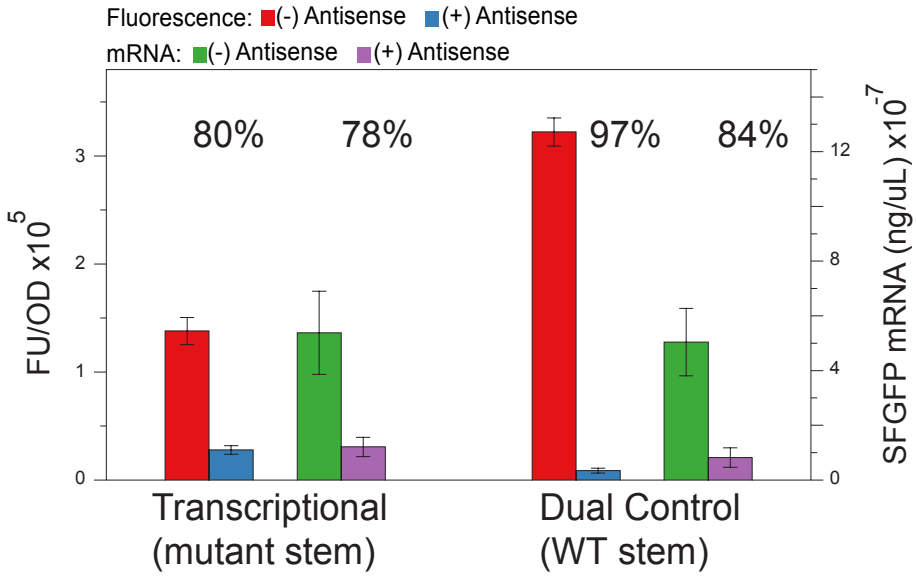
causing GFP to be expressed. Bulk fluorescence was measured using a plate reader (see Materials and Methods). The no theophylline condition is shown in green for the transcriptional cascade and blue for dual control. The dual control reduces circuit leak and the background fluorescence. The transformation, culturing, and measurements were done on different days for three days of experiments. The colored region indicates the standard deviation from three biological replicates.



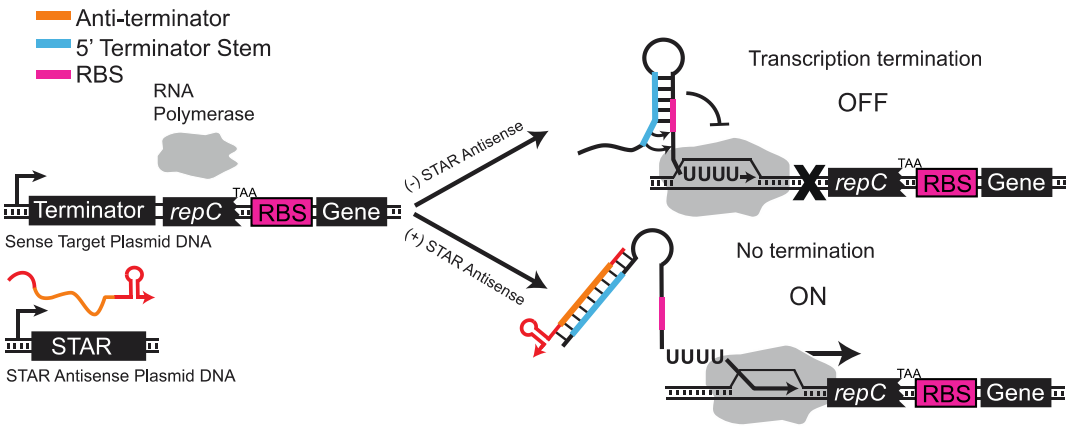
Supplementary Figure A3: Functional data for the transcriptional and dual control repressor cascades over time performed from glycerol stocked strains. Experiment performed as described in Supplementary Note A1. Data was analyzed as described in the Materials and Methods. The no theophylline condition is shown in green for the transcriptional cascade and blue for dual control. The dual control reduces circuit leak and the background fluorescence. The colored region indicates the standard deviation from three biological replicates.



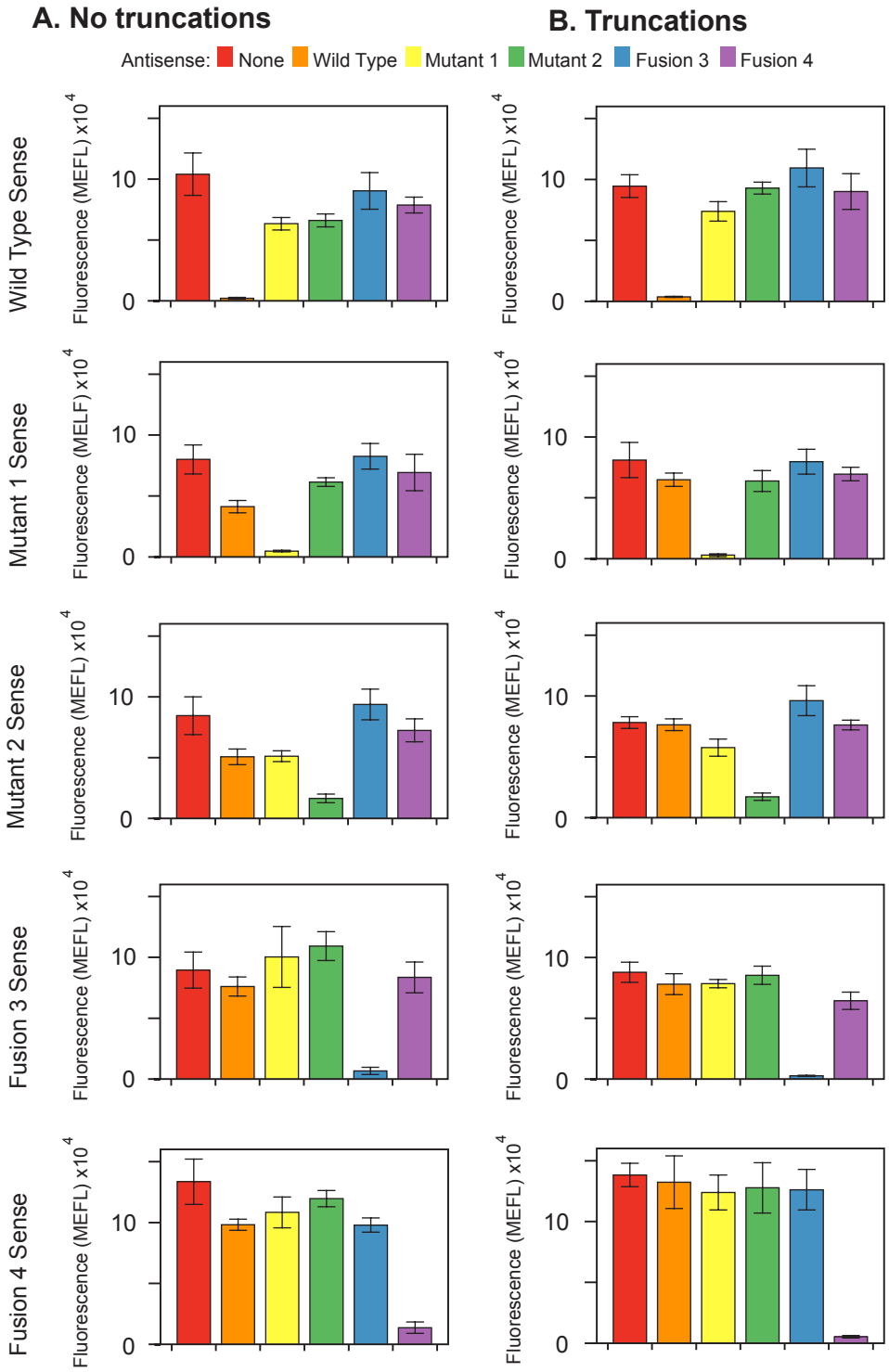
Supplementary Figure A4: Functionality of the 12nt RepC fusion versus complete RepC knockout. (A) Schematic of the pT181 transcriptional terminator with RBS (pink) and dual control 12nt fusion to RepC (orange) expressing SFGFP (green). The dual control with no RepC lacks the orange region. (B) Experiments were performed as described in Supplementary Note A1. Bulk fluorescence data was collected using a plate reader. The 12nt RepC fusion shows similar percent repression (97% +/- 1.5%) as the dual control with no RepC (96% +/-0.8%), but has a much higher ON level. Error bars represent the standard deviation of nine biological replicates.



Supplementary Figure A5: Quantitative RT-PCR of a transcriptional and a dual control pT181 attenuator. qRT-PCR quantification was performed as described in Supplementary Note A1. Bulk fluorescence was measured using a plate reader from the same cultures as the RNA was extracted. The fluorescence ON levels (red) were significantly different ($P > 0.05$ for Welch's t-test), but mRNA ON levels (green) were not significantly different. Transcriptional fluorescence repression (80% +/- 3.4%) was similar to mRNA repression (78 +/- 9.1%) while dual control fluorescence repression (97% +/- 0.7%) was much higher than mRNA repression (84% +/- 8.1%). The error bars represent the standard deviation of three technical replicates of three biological replicates.

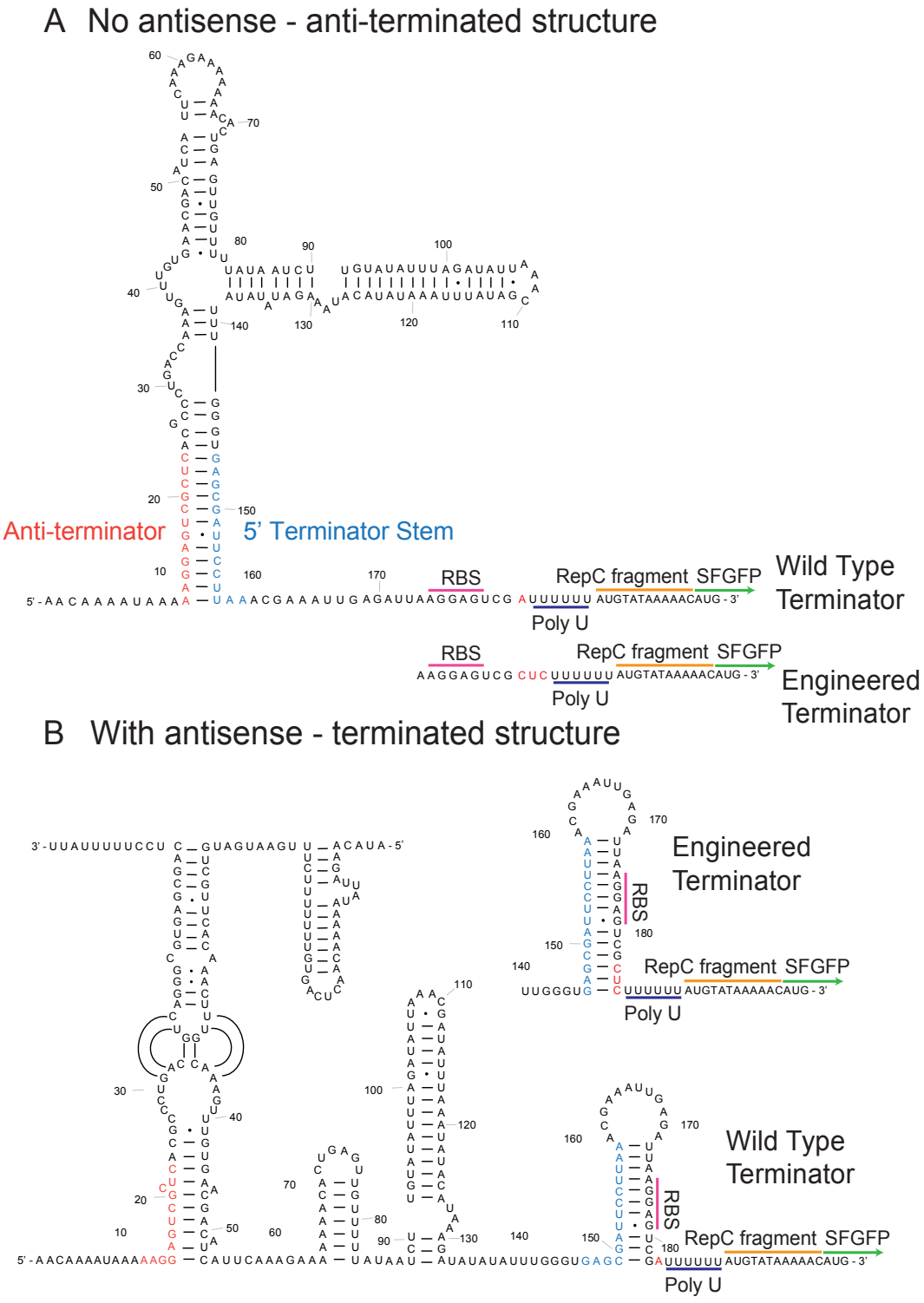


Supplementary Figure A6: Schematic of the pT181 anti-terminator STAR mechanism. The sense target region consists of the pT181 STAR target region from Chappell et al. (2) followed by a 96 nt fragment of the *repC* gene ending in a stop codon, TAA, included as a transcriptional fusion before a ribosome binding site (RBS) and the regulated gene of interest. In the absence of the STAR RNA (red/orange), the terminator forms, preventing downstream transcription by RNA polymerase (grey). Thus in the absence of STAR RNA the mechanism is transcriptionally OFF. The STAR RNA contains an anti-terminator sequence (orange) complementary to the 5' half of the terminator (blue). When present, the STAR RNA binds to the terminator, preventing terminator formation and allowing transcription elongation. Thus in the presence of STAR RNA the mechanism is transcriptionally ON.

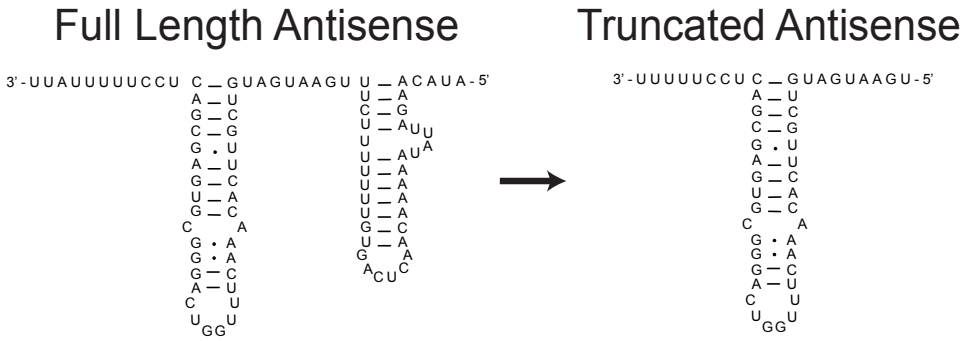


Supplementary Figure A7: In vivo expression data used to calculate orthogonality matrices in Figure 2.5. Functional characterization of orthogonality matrix cognate and non-cognate pairs. Average fluorescence (MEFL) was collected by flow cytometry of *E. coli* TG1 cells with no antisense (red), wild type antisense (orange),

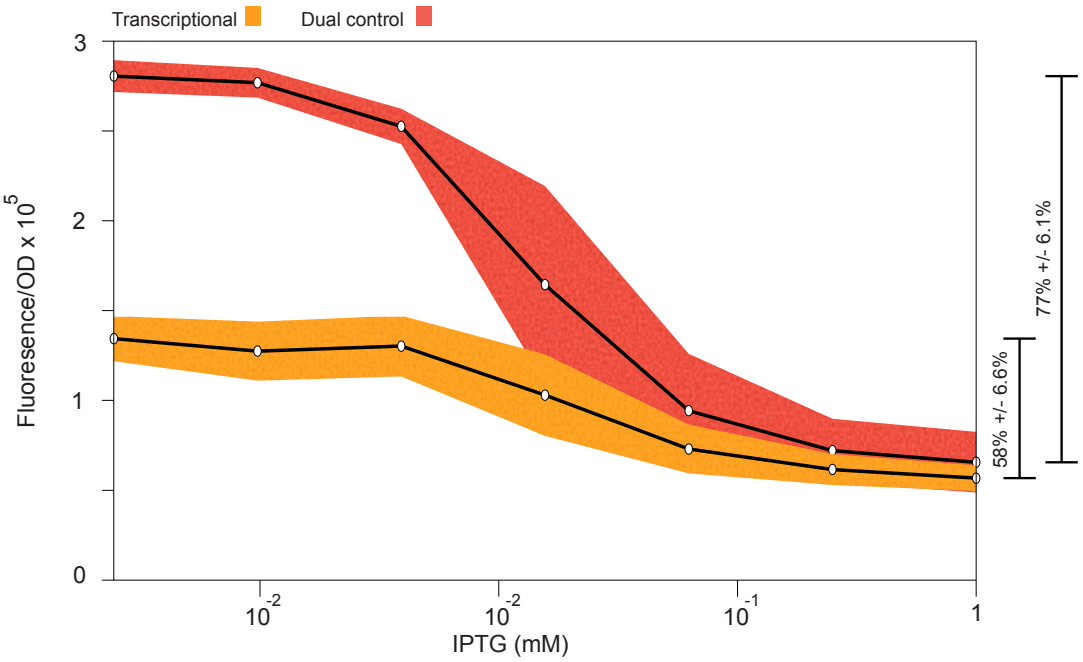
mutant 1 antisense (yellow), mutant 2 antisense (green), fusion 3 antisense (blue), or fusion 4 antisense (purple). Error bars represent standard deviations of at least seven biological replicates.



Supplementary Figure A8: Sequences and model structures of the schematic shown in Figure 2.1B with both the anti-terminated structure (ON) in the absence of antisense (A) and the terminated structure (OFF) in the presence of antisense (B). The anti-terminator region is highlighted in orange and the 5' terminator stem is highlighted in blue. The RBS is shown with a pink line, the Poly U with a blue line, the 12 nt RepC fragment with an orange line, and SFGFP coding sequence with a green line. Sequences for both the wild type and engineered terminators are shown. Structures are from Brantl and Wagner (3).



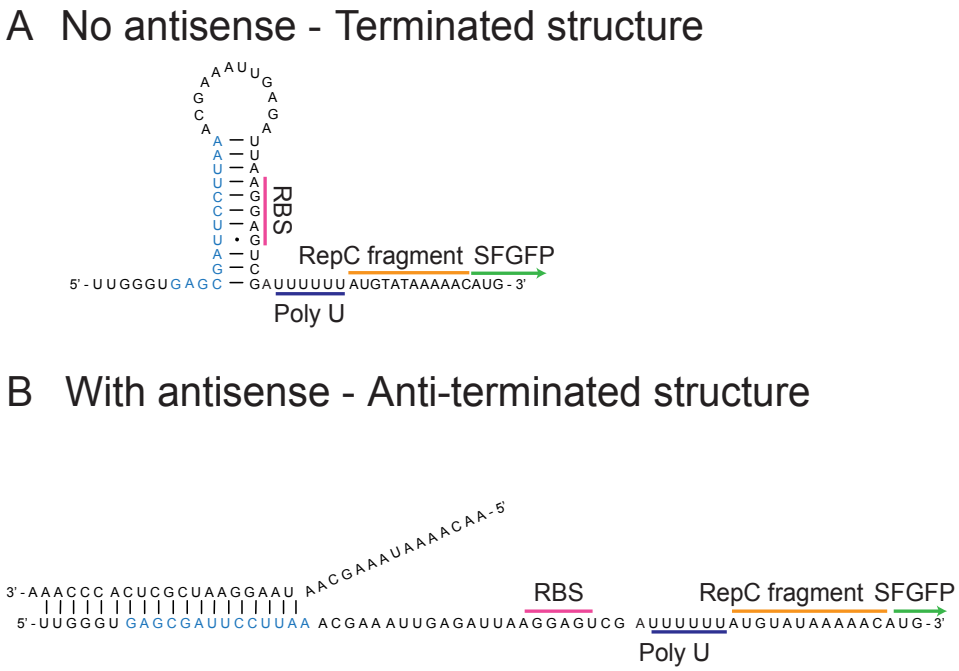
Supplementary Figure A9: Sequence and structure of the full length wild type pT181 antisense and the truncated wild type antisense. In order to create the truncated antisense, the first hairpin from the 5' side was removed and sequence was truncated from the 5' stem of the antisense. Structures are from Brantl and Wagner (3).



Supplementary Figure A10: Induction curves of transcriptional and dual control attenuators. Antisense expression is under the control of the Lac promoter. Experiment performed as described in Supplementary Note A1. Data was analyzed as described in Materials and Methods. The dual control attenuator (red) shows greater dynamic range (77% +/- 6.1%) in response to IPTG induction than the transcriptional (orange) attenuator (58% +/- 6.6%). The colored region indicates the standard deviation from 3 biological replicates.

Supplementary Table A3: Averages and standard deviations (MEFL) plotted in Figure 2.2A and 2.2B. Values have been rounded to the nearest integer for ease of interpretation.

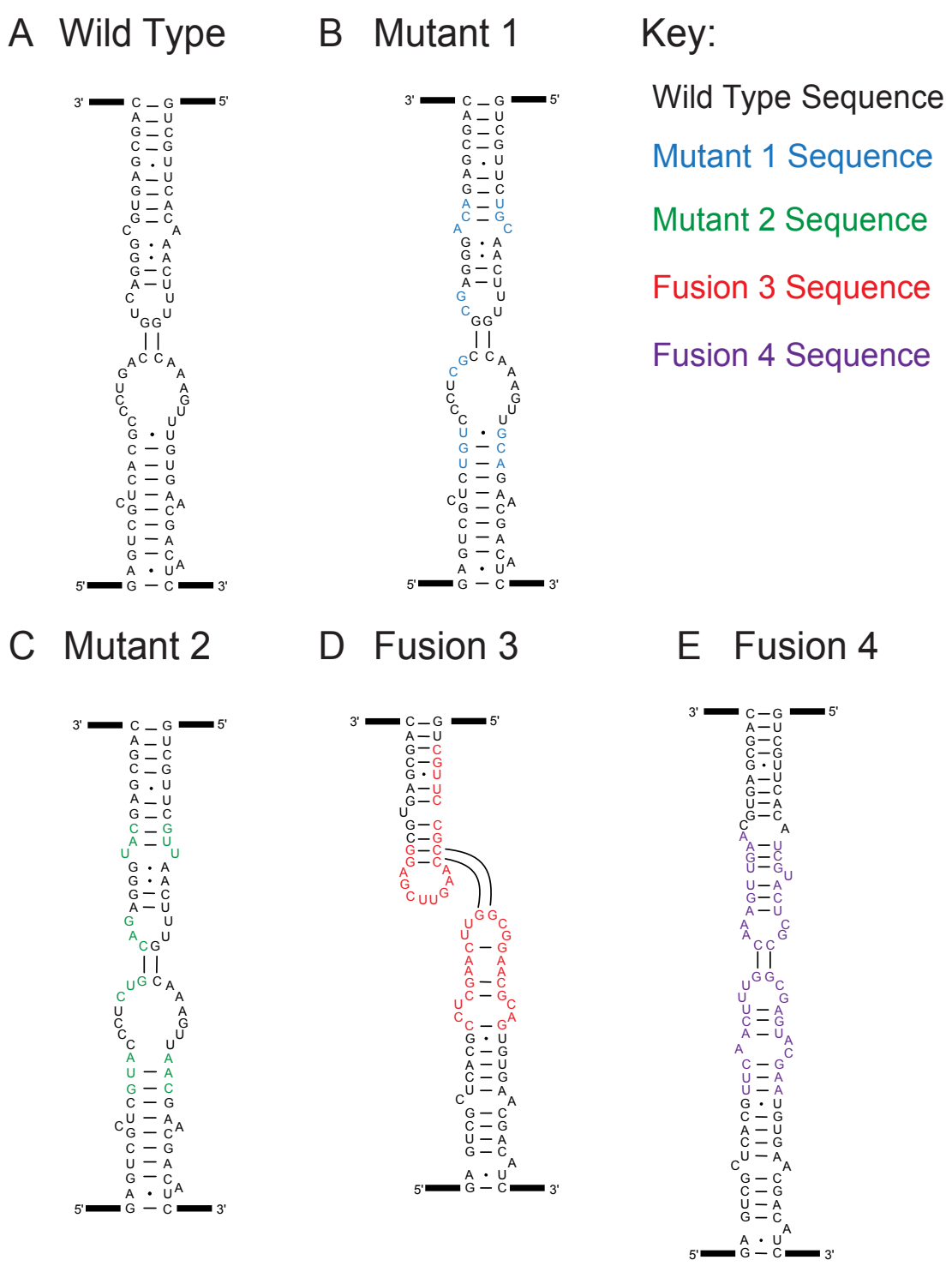
	No antisense (average)	With antisense (average)	No antisense (SD)	With antisense (SD)
Transcriptional WT terminator	40289	14857	7307	1659
Dual control WT terminator	94148	1630	11530	260
Transcriptional Eng terminator	38699	5849	4207	1132
Dual control Eng terminator	27401	671	3644	181



Supplementary Figure A11: Sequences and structures of the schematic shown in Figure 2.3A with both the terminated structure (OFF) in the absence of antisense (A) and the anti-terminated structure (ON) in the presence of antisense (B). The 5' terminator stem is highlighted in blue. The RBS is shown with a pink line, the Poly U with a blue line, the 12 nt RepC fragment with an orange line, and SFGFP with a green line. Structures are derived from diagrams in Chappell et al. (2).

Supplementary Table A4: Averages and standard deviations (MEFL) plotted in Figure 2.3B and 2.3C. Values have been rounded to the nearest integer for ease of interpretation.

	No antisense (average)	With antisense (average)	No antisense (SD)	With antisense (SD)
STAR	3700	37976	695	11837
Dual control STAR	115	106062	16	19872



Supplementary Figure A12: Sequences and structures of antisense/attenuator interaction regions in Figure 2.4A including the wild type (A), mutant 1 (B), mutant 2 (C), fusion 3 (D), and fusion 4 (E). Wild type sequence is shown in black. Mutant 1 sequence is shown in blue. Mutant 2 sequence is shown in green. Fusion 3 sequence

is shown in red. Fusion 4 sequence is shown in purple. Structures are from Brantl and Wagner (3) (A), Lucks et al. (4) (B, C) and Takahashi et al. (5) (D, E).

Supplementary Table A5: Averages and standard deviations (MEFL) plotted in Figure 2.4B and 2.4C. Values have been rounded to the nearest integer for ease of interpretation.

	No antisense (average)	With antisense (average)	No antisense (SD)	With antisense (SD)
Transcriptional WT	38699	5849	4207	1132
Dual control WT	94590	3671	9334	455
Transcriptional Mutant 1	38311	7369	4557	776
Dual control Mutant 1	81071	2832	14435	1116
Transcriptional Mutant 2	37943	14133	2518	867
Dual control Mutant 2	78197	17278	4814	3053
Transcriptional Fusion 3	36984	6935	3957	844
Dual control Fusion 3	87808	2675	8282	432
Transcriptional Fusion 4	48829	10917	4764	2190
Dual control Fusion 4	138298	5338	9637	1056

Supplementary Table A6: Standard deviations for the percent repressions presented in Figure 2.5B with full-length antisense. Bolded cells indicate the diagonal.

	WT full AS	Mut 1 full AS	Mut 2 full AS	Fus 3 full AS	Fus 4 full AS
WT sense	**0.77**	11.31	11.82	20.47	14.14
Mut 1 sense	9.89	**1.23**	12.25	20.17	22.70
Mut 2 sense	13.39	12.32	**5.50**	25.33	19.28
Fus 3 sense	16.57	33.62	24.14	**3.56**	20.95
Fus 4 sense	10.78	14.73	13.43	11.11	**3.78**

Supplementary Table A7: Standard deviations for the percent repression presented in Figure 2.5C with truncated antisense. Bolded cells indicate the diagonal.

	WT trunc AS	Mut 1 trunc AS	Mut 2 trunc AS	Fus 3 trunc AS	Fus 4 trunc AS
WT sense	**0.62**	11.42	10.99	19.93	18.07
Mut 1 sense	15.76	**1.51**	17.57	21.57	16.72
Mut 2 sense	8.61	10.06	**4.13**	17.50	7.83
Fus 3 sense	12.92	9.30	12.55	**0.57**	10.65
Fus 4 sense	17.02	12.14	16.28	13.62	**0.81**

A.2 References

1. Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., et al. (2009) The MIQE Guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.*, **55**, 611-622.

2. Chappell,J., Takahashi,M.K. and Lucks,J.B. (2015) Creating small transcription activating RNAs. *Nat. Chem. Biol.*, **11**, 214–220.

3. Brantl,S. and Wagner,E.G. (2000) Antisense RNA-mediated transcriptional attenuation: an in vitro study of plasmid pT181. *Mol. Microbiol.*, **35**, 1469–1482.

4. Lucks,J.B., Qi,L., Mutalik,V.K., Wang,D. and Arkin,A.P. (2011) Versatile RNA-sensing transcriptional regulators for engineering genetic networks. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 8617–8622.

5. Takahashi,M.K. and Lucks,J.B. (2013) A modular strategy for engineering orthogonal chimeric RNA transcription regulators. *Nucleic Acids Research*, **41**, 7577–7588.

APPENDIX B

Supplementary information for Timescale Differences Allow Construction of an RNA-based Incoherent Feed Forward Loop

B.1 Supplementary Figures and Tables

	Description
Table B1	Parameter values used in Figure 4.3 and 4.5
Table B2	Important DNA sequences
Table B3	Plasmids used in this study
Figure B1	Complete CRISPRi pre-incubation data
Figure B2	Complete STAR parameterization data
Figure B3	Complete CRISPRi parameterization data
Figure B4	Parameter correlation comparison of the 1000 fitted STAR parameters
Figure B5	Parameter correlation comparison of the 1000 fitted CRISPRi parameters
Figure B6	Predictions of all combinations of STAR and CRISPRi parameter sets
Figure B7	Complete pulse generator circuit data
Figure B8	Prediction parameters compared to fitting parameters
Figure B9	Best fitting from each of the 10 Bayesian inference for STAR
Note B1	Slope computation and fitting for Figure 4.2

Table B1. Parameter values used in Figure 3 and 5. The best fitting EGFP reporter parameters from both STAR and CIRSPRi pre-incubation experiments are listed in order of STAR/CRISPRi.

Parameters	Varying Interval	Best Fitting Values	Best Prediction Values
α_s	10^{-2} - 10 s^{-1}	2.038 s^{-1}	2.038 s^{-1}
δ_s	10^{-5} - 10^{-1} s^{-1}	0.0915 s^{-1}	0.0915 s^{-1}
β_s	10^3 - 10^7 $nM^{-1}s^{-1}$	310969 $nM^{-1}s^{-1}$	310969 $nM^{-1}s^{-1}$
$\alpha_{cr}/_{tr}$	10^{-2} – 10 s^{-1}	0.739 s^{-1}	0.03 s^{-1}
δ_{cr}	10^{-5} - 10^{-1} s^{-1}	0.0615 s^{-1}	0.0614 s^{-1}
δ_{tr}	10^{-5} - 10^{-1} s^{-1}	0.030 s^{-1}	0.073 s^{-1}
δ_g	10^{-5} - 10^{-1} s^{-1}	0.085 s^{-1}	0.0176 s^{-1}
γ_1	10^3 - 10^7 $nM^{-1}s^{-1}$	7.38x10^6 $nM^{-1}s^{-1}$	8.69x10^6 $nM^{-1}s^{-1}$
γ_2	10^3 - 10^7 $nM^{-1}s^{-1}$	6.56x10^6 $nM^{-1}s^{-1}$	3.68x10^6 $nM^{-1}s^{-1}$
ω	10^3 - 10^7 $nM^{-1}s^{-1}$	3.1x10^4 $nM^{-1}s^{-1}$	7.17x10^6 $nM^{-1}s^{-1}$
α_m	10^{-1} – 100 s^{-1}	0.1/50.65 s^{-1}	0.1 s^{-1}
δ_m	10^{-5} - 10^{-1} s^{-1}	4.01x10^{-4}/0.094 s^{-1}	4.01x10^{-4} s^{-1}
K_i	10^{-4} - 10^{-2} s^{-1}	0.0012/9.019x10^{-4} s^{-1}	0.0012 s^{-1}
K_e	10^{-4} - 10^{-2} s^{-1}	0.009/4.168x10^{-4} s^{-1}	0.009 s^{-1}
α_{gm}	10^{-3} - 10^{-1} s^{-1}	0.092/0.015 s^{-1}	0.092 s^{-1}

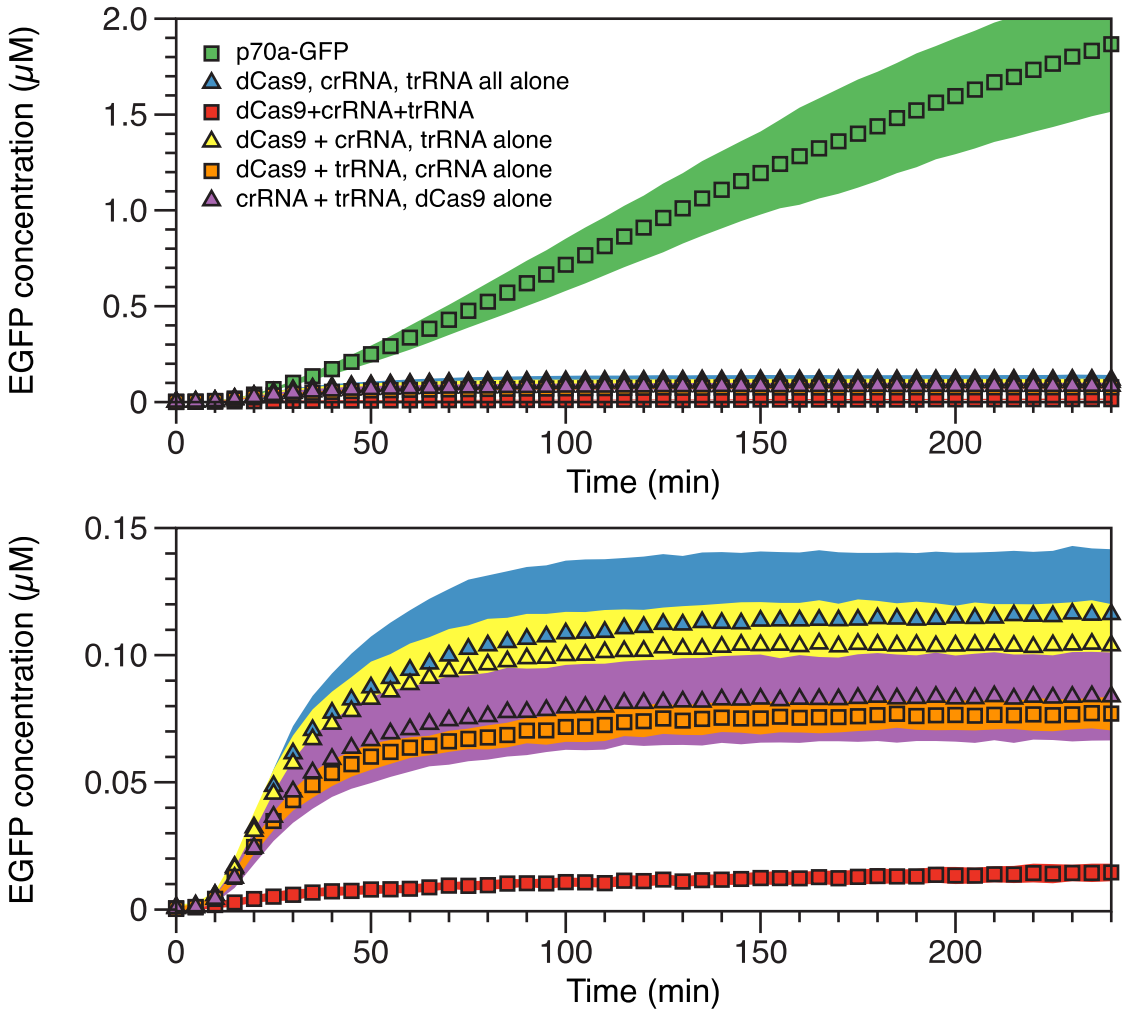
Table B2. Important DNA sequences

Name	Sequence
J23119	TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT
p70a	TGAGCTAACACCGTGCGTGTTGACAATTTTACCTCTGGCGGTGATAATGGTTGCA
STAR 5	TGAACTGTATACATTCCCCGCAGGATAGGAATTGAAGATGAAACGATGAGACTTGGGACGAGGATCT
STAR 5 Target	TCGTCCCAAGTCTCATCGTTTCATCTTCAATTCCTATCCTGCGGGGAATGTATACAGTTCATGTATATATTCCCCGCTTTTTTTTTGGATCT
crRNA	GGTAAAATTGTCAACACGCAGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAAC
Scrambled crRNA	AAGCAGATTACGTTCAAGCAGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAAC
tracrRNA	ATCTTGTTGGAACCATTCAAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTG
eGFP (Ribosome binding site (RBS) - GFP)	AGAAGGAGATATACCATGGAGCTTTTCACTGGCGTTGTTCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACA

	TCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCA TGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACA ACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCC CCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCT GGAGTTCGTGACCGCCGCCGGGATCTAACTCGAG
TrrnB	GAAGCTTGGGCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGC CGTCGACCATCATCATCATCATCATTGAGTTTAAACGGTCTCCAGCTTGGC TGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGA ACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGT GGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGA TGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAA TAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTT TGTCGGTGAACT
T500	CAAAGCCCGCCGAAAGGCGGGCTTTT

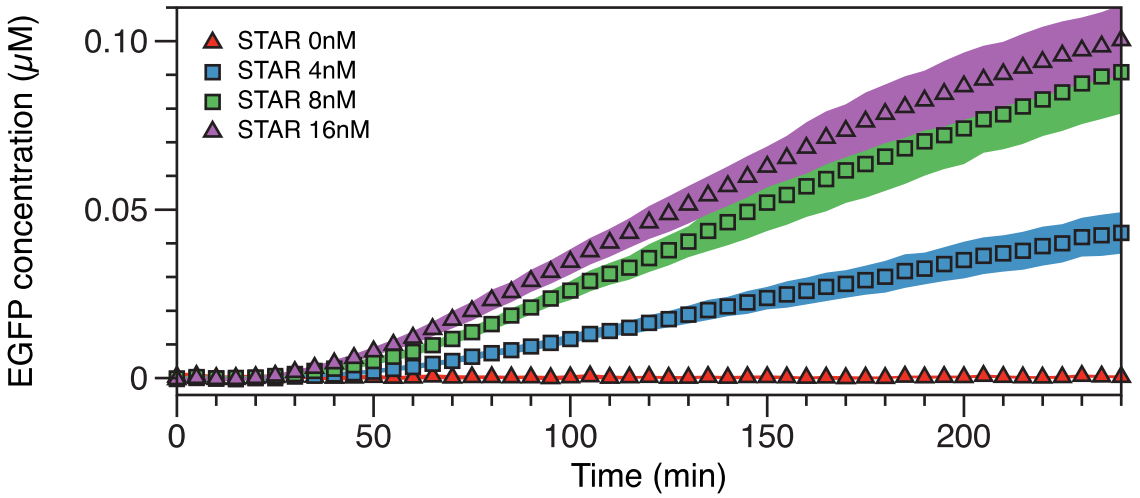
Table B3. Plasmids used in this study. Sequences in the plasmid architecture can be found in Table B2.

Plasmid #	Plasmid architecture	Name	Figure	Reference
JBL002	J23119 – TrrnB – ColE1 origin – AmpR	No STAR control	2-6, S2-5, S7, S10	A
70a-GFP	p70a – GFP – ColE1 origin - AmpR	P70a-GFP	2	Garamella et al. 2016
AMW019	p70a – STAR 5 Target – GFP – ColE1 origin - AmpR	P70a-STAR Target-GFP	2, 4, 5, S4, S5, S7	This paper
JBL4971	J23119 – STAR 5 – t500 – ColE1 origin - AmpR	STAR 5	2, 4, S5, S10	Chappell et al. 2017
CSM257	J23119 – crRNA – t500 – CamR – ColE1	crRNA	4	This paper
CSM258	J23119 – scrambled crRNA – t500 – CamR – ColE1	crRNA control	4, 5, S7	This paper
CSM275	J23119 – tracrRNA – t500 – CamR – ColE1	tracrRNA	4	This paper

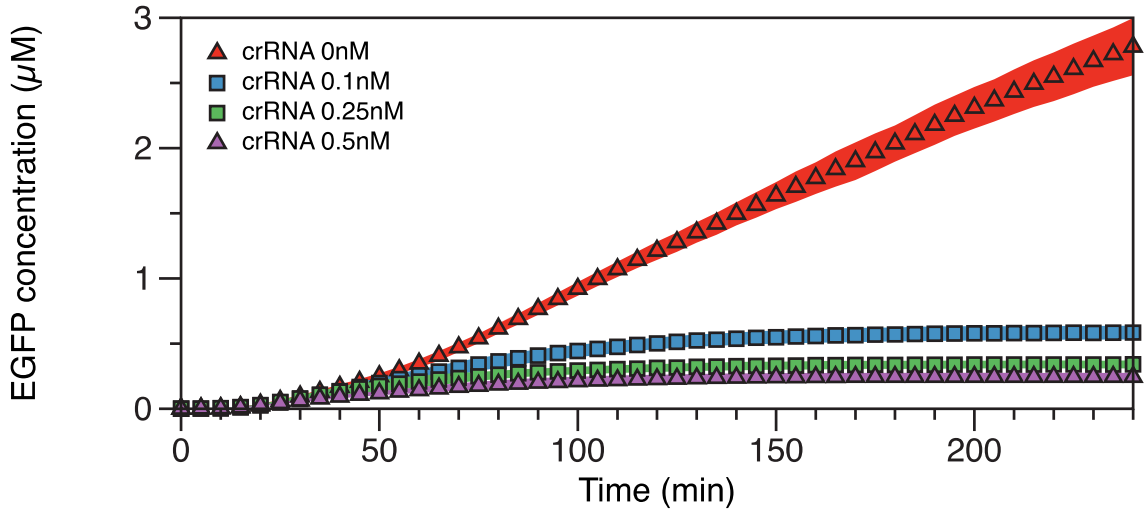


Supplementary Figure B1. Complete data for the CRISPRi pre-incubation experiment. Functional time course characterization of the CRISPRi response when parts are incubated together or alone in various combinations for 2 hours before measurements. When no crRNA or trRNA is present, GFP expression is ON (green). When all the parts are incubated together, GFP is quickly repressed (red). GFP repression is delayed when all parts are incubated separately (blue). Incubating trRNA alone, but dCas9 and crRNA together shows similar delays in repression (yellow). When crRNA (orange) or dCas9 (purple) is incubated alone (but dCas9 and trRNA together or crRNA and trRNA together, respectively) the delay is less dramatic. The colored region indicates the standard deviation of nine replicates. The same data

is shown in both plots except the p70a-GFP has been removed from the bottom plot so that the other conditions could be seen clearly.

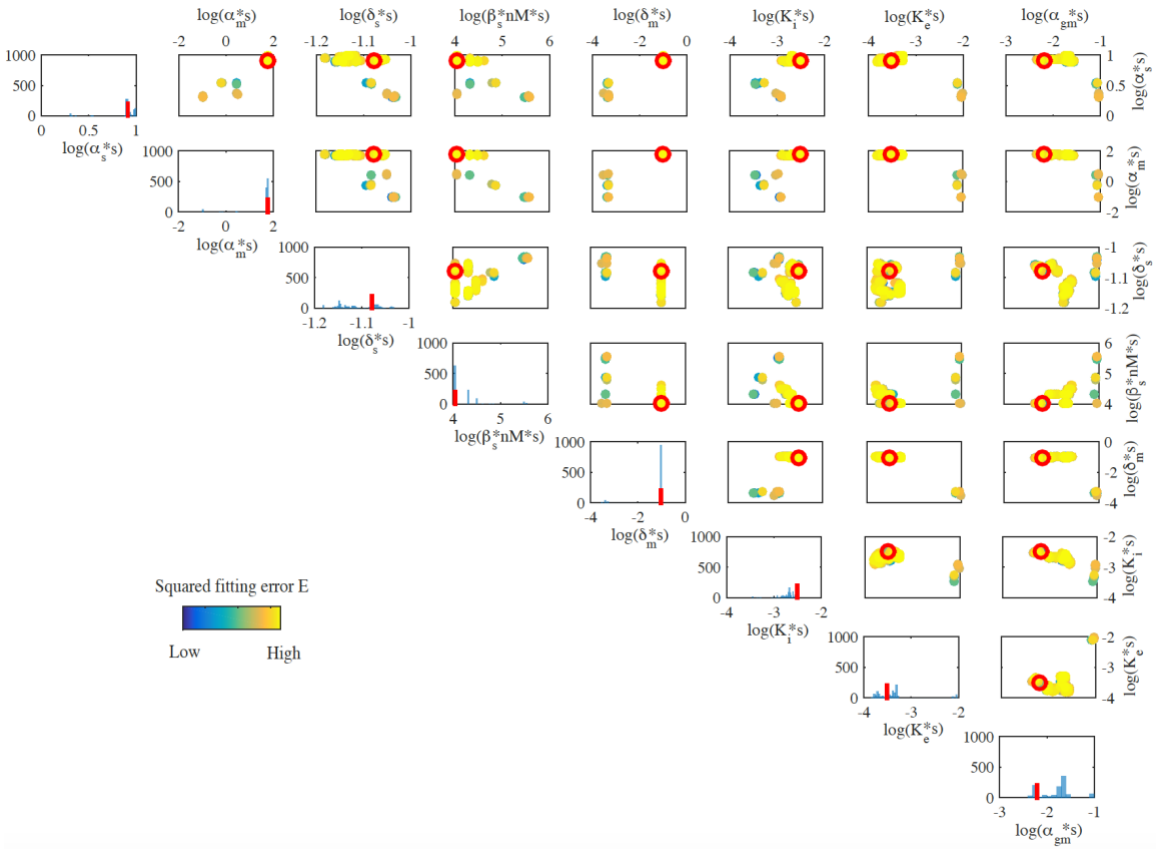


Supplementary Figure B2. Complete data for the STAR parameterization experiment. Functional time course characterization of GFP expression when different concentrations of STAR plasmid is added to the TXTL reaction with 0.5nM of the p70a-STAR Target-GFP plasmid at the start of the measurement. The colored region indicates the standard deviation of nine replicates.

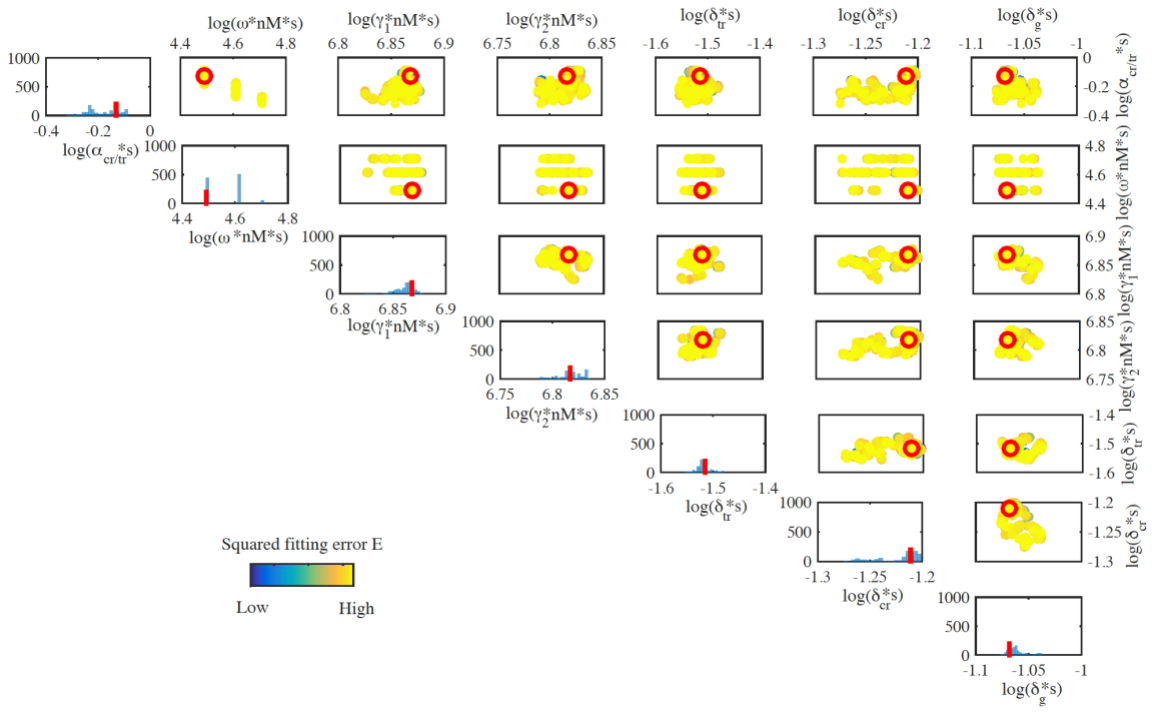


Supplementary Figure B3. Complete data for the CRISPRi parameterization experiment. Functional time course characterization of GFP expression when different concentrations of crRNA and trRNA plasmids are added to the dCas9 TXTL

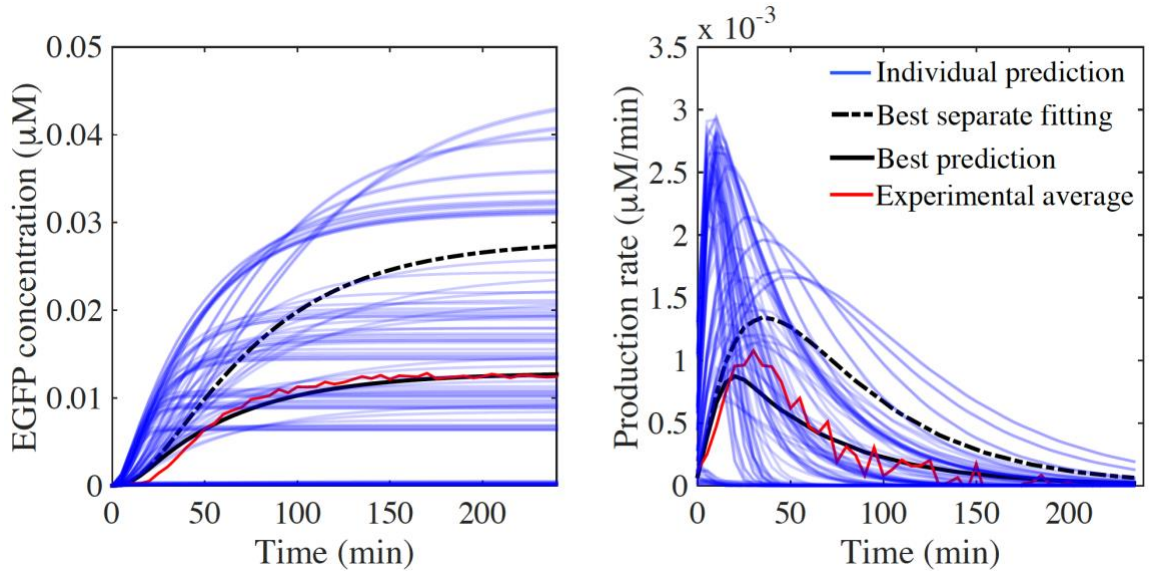
reaction with 0.5nM of the p70a-GFP plasmid at the start of the measurement. The colored region indicates the standard deviation of nine replicates.



Supplementary Figure B4. Distribution and parameter correlation comparison of the 1000 fitted STAR parameters that gave the lowest fitting error to the STAR experiments. While no strong correlations were observed among the parameters, some parameters have a wider distribution such as β_s and K_i, some others have narrower distributions such as α_m.

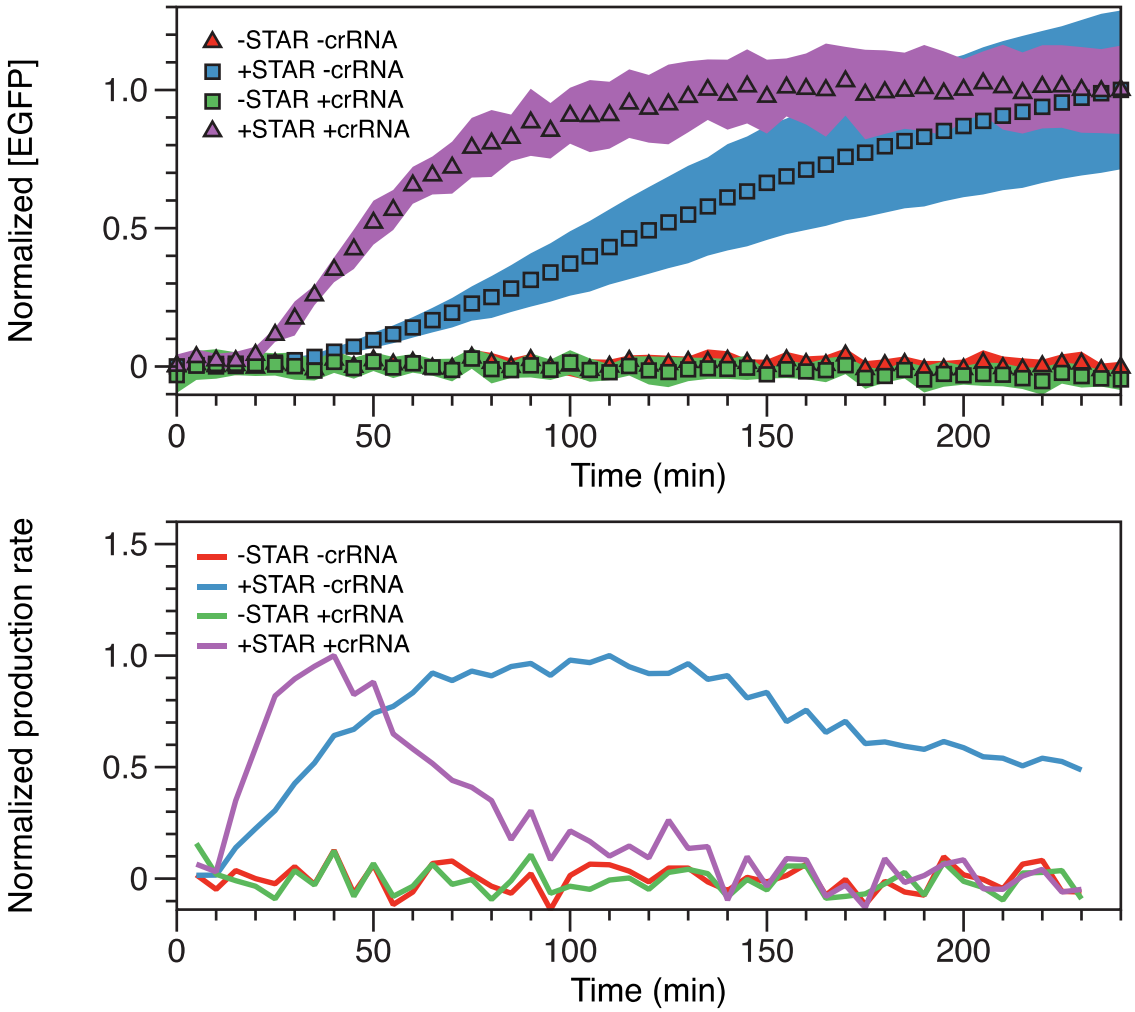


Supplementary Figure B5. Distribution and parameter correlation of the 1000 fitted CRISPRi parameters that gave the lowest fitting error to the CRISPRi experiments. While minimal to no correlations were observed among the parameters, some parameters have a wider distribution such as γ_1, some others have narrower distributions such as ω.



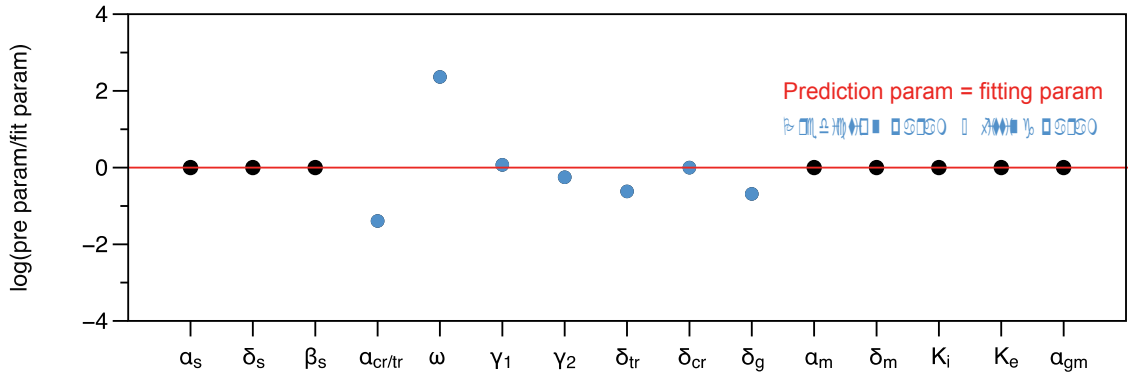
Supplementary Figure B6. Predictions of all the 100 combinations of STAR and CRISPRi separately fitted parameter sets demonstrate pulse in the combined model.

Solid blue plots are individual predictions, dashed black plot is the prediction with the best separately fitted parameters, the solid black plot is the best prediction out of this 100 combinations, and the solid red plot is the averaged experimental measurement.

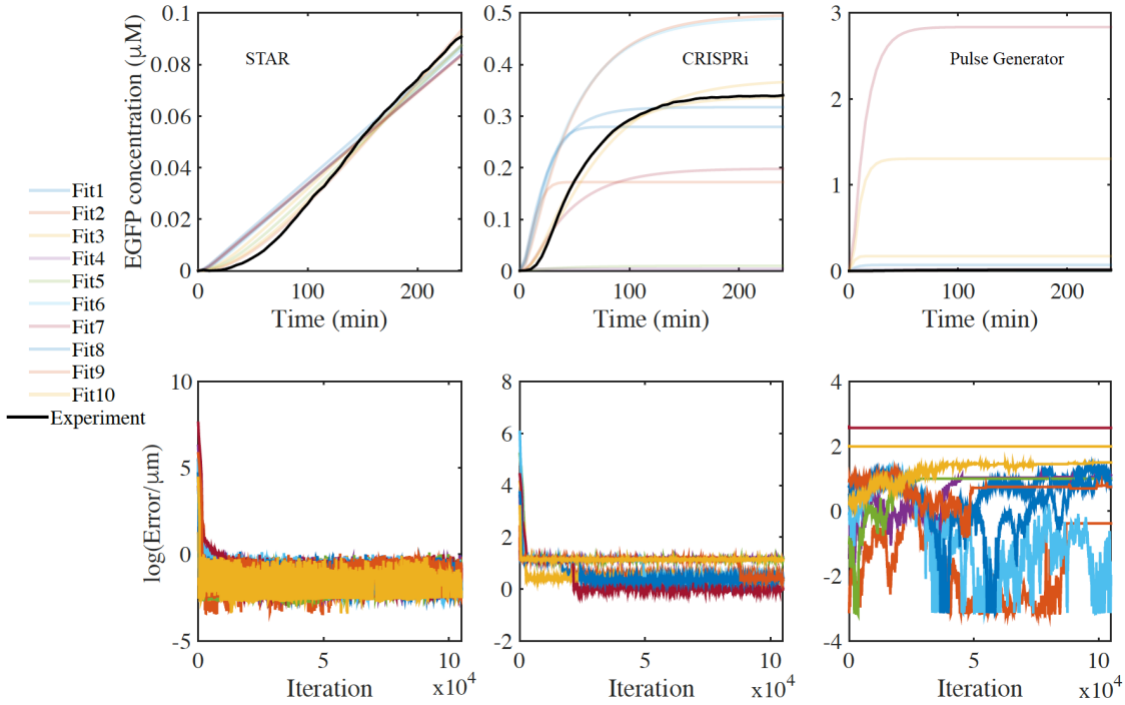


Supplementary Figure B7. Complete data for the pulse generator experiment. Functional time course characterization of pulse generator with 0.5nM of the p70a-STAR Target-GFP plasmid, 16nM of STAR plasmid (for the +STAR conditions), and 0.25nM of crRNA and trRNA plasmids (for the +crRNA conditions). The pulse generator activates expression and plateaus quickly. A pulse is generated in production rate when both STAR and crRNA are present (+STAR +crRNA, purple). We show both the normalized concentrations (top) and normalized production rate

(bottom) because GFP does not degrade so the pulse is seen in production rate. The production rates shown are smoothed averages of nine replicates. The pulse generator (+STAR +crRNA, purple), negative control (-STAR -crRNA, red), and CRISPRi control (-STAR +crRNA, green) conditions are normalized by the +STAR +crRNA and the STAR control (+STAR -crRNA, blue) is normalized by its own maximum. The colored regions (top) indicate the standard deviation of nine replicates.



Supplementary Figure B8. Comparison in terms of parameter values in the best fitted and prediction parameter sets, demonstrating that a faster CRISPRi repressor formation rate was needed in the pulse circuit model for better prediction accuracy.



Supplementary Figure B9. Best fitting from each of the 10 Bayesian inference for STAR (a), CRISPRi (b), and pulse generator model (c). Fitting error demonstrates convergence comparison in log scale for STAR (e), CRISPRi (f), and pulse generator (h). There are 8, 12, and 15 parameters fitted in the STAR, CRISPRi, and pulse generator model respectively. Due to the increased number of parameters, even with a larger number of iterations it was challenging to obtain convergence for the CRISPRi and pulse model fitting, with worse fitting performance with respect to to STAR.This indicates that our fitting approach becomes more challenging and computationally expensive as the number of fitted parameters increases.

Supplementary Note B1: Slope computation and fitting for Figure 4.2 in the main manuscript.

Data were first smoothed by employing a 2 point moving average filter with MATLAB. The slope of each kinetic data set was computed using the MATLAB function diff. The

slope was then averaged over all the available samples. The averaged slope was fitted to the following functions, using MATLAB's lsqcurvefit routine:

a) STAR system:

$$s(t) = d_1(1 - e^{-\frac{t}{d_2}})$$

Fitted values for the parameters are: d1 = 0.3*10-3 μm/min, d2 = 25 min.

b) CRISPR system:

$$c(t) = k_1(e^{-\frac{t}{k_2}}) + k_3(e^{-\frac{t}{k_4}})$$

Fitted values for the parameters are: k1 = -0.0218 μm/min, k2 = 14.24 min, k3 = 0.0224 μm/min, k4 = 18.27 min

APPENDIX C

Supplementary Information for Chapter 5

C.1 Supplementary Tables

Table C1: Model Variables

Variable	Description
t	time
I_g	Intensity of green light
I_r	Intensity of red light
A	Concentration of repressor A
B	Concentration of repressor B
C	Concentration of repressor C
D	Concentration of repressor D
E	Concentration of repressor E
F	Concentration of repressor F
G	Concentration of repressor G
M_g	Concentration of GFP mRNA
M_r	Concentration of RFP mRNA
P_g	Concentration of GFP
P_r	Concentration of RFP

Table C2: Parameter estimations from the literature for both models.

Name	Value	Description	Reference
$k_{p,g}(I_g)$	varies	Transcription rate based on CcaS/CcaR model	Olson et al. 2014
$k_{p,r}(I_r)$	varies	Transcription rate based on Cph8/OmpR model	Olson et al. 2014
$c(t-t_{delay})$	varies	Set point based on CcaS/CcaR or Cph8/OmpR model	Olson et al. 2014
β	230 molecules/ min	RNA transcription constant	Hu and Lucks, 2015
K	150, 120 molecules	Repression constant	Hu and Lucks, 2015
d_m	0.23/min	RNA degradation constant	Estimated from bionumbers
k_{Pg}, k_{Pr}	1320/min	Protein translation rate	Estimated from bionumbers
d_{Pg}, d_{Pr}	0.023/min	Protein degradation rate	Estimated from bionumbers
$\eta CD, \eta DC, \eta AE, \eta GE, \eta GF, \eta EG$	2	Number of tandem attenuators (ηXW) for X attenuator regulating W antisense	Determined from stability analysis

$\eta AC, \eta BD,$ $\eta GD, \eta DF,$ $\eta BG, \eta FG$	1	Number of tandem attenuators (ηXW) for X attenuator regulating W antisense	Determined from stability analysis